Structure maximale pour la somme des carrés d'une contingence aux marges fixées: une solution algorithmique programmée

Israël-César Lerman, Philippe Peter

UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France

Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 608

STRUCTURE MAXIMALE POUR LA SOMME DES CARRÉS D'UNE CONTINGENCE AUX MARGES FIXÉES : UNE SOLUTION ALGORITHMIQUE PROGRAMMÉE

Israël-César LERMAN
Philippe PETER

Février 1987

IRISA INSTITUT DE RECHERCHE EN INFORMATIQUE
ET SYSTÈMES ALÉATOIRES

Campus Universitaire de Beaulieu
Avenue du Général Leclerc
35042 - RENNES CÉDEX
FRANCE
Tél. : (99) 36.20.00
Télex : UNIRISA 95 0473 F

STRUCTURE MAXIMALE POUR LA SOMME DES CARRES D'UNE CONTINGENCE AUX MARGES FIXEES ; UNE SOLUTION ALGORITHMIQUE PROGRAMMEE

Israël-César **LERMAN**
Philippe **PETER**

Publication Interne n° **318** - Octobre 1986 - 90 pages

RESUME :

Ce problème est reconnu comme très difficile voire même quasi impossible à résoudre. Nous commençons par analyser l'approche -initialisée par différents auteurs- que nous appelons "analytique" où une borne majorant la borne exacte est fournie par une formule mathématique, symétrique par rapport aux deux marges. Nous montrons que cette approche est nécessairement basée sur une application de l'inégalité de Shwartz, mais conçue dans un cadre logique. Nous déterminons la meilleure borne pouvant être trouvée dans ce cadre et montrons que cette dernière est médiocre par rapport à celle, dissymétrique, obtenue à partir d'une seule marge.
La solution originale que nous proposons pour construire la configuration optimale et la borne exacte associée, consiste d'abord à remplacer la notion de formule par celle beaucoup plus générale d'algorithme récursif dont le point de départ est le couple des marges du tableau de contingence. La majorité de ce rapport est réservée à l'étude de cette solution. Des notions spécifiques sont introduites et la justification mathématique de l'algorithme est fournie de la façon la plus élaborée.

MAXIMAL ASSOCIATION STRUCTURE FOR THE SUM OF SQUARES OF A CONTINGENCY TABLE ENTRIES : A PROGRAMMED ALGORITHMIC SOLUTION

ABSTRACT :

This problem is considered as very difficult -and even quasi-impossible- to resolve. We begin by analyzing the approach that we call "analytic" where a non exact bound is given by a mathematic formulae, this last being symmetric with respect to the two margins. We show that this approach is necessarily based on an application of Shwartz inequality, conceived in a logical context.

.../...

The best bound which can be obtained by this method is in fact less accurate than this one determined from the "best" margin.
The original solution that we propose to construct the optimal configuration and the associated exact bound, is based on the notion of recursive algorithm. The starting state of application of the algorithm is the couple of margins of the contingency table. Most important part of this report is devoted to the study of this solution. Specific notions are introduced. On the other hand, mathematical justification of the algorithm is provided as deeply as possible.

I. INTRODUCTION ; POSITION DU PROBLEME

	1	...	j	...	J	
1	c_{11}		c_{1j}		c_{1J}	a_1
⋮						
i	c_{i1}		c_{ij}		c_{iJ}	a_i
⋮						
I	c_{I1}		c_{Ij}		c_{IJ}	a_I
	b_1		b_j		b_J	n

Table 1

On désigne par $\{c_{ij}/1 \leqslant i \leqslant I,\ 1 \leqslant j \leqslant J\}$ une table de contingence à I lignes et J colonnes. On désigne par $\{a_i/1 \leqslant i \leqslant I\}$ la marge colonne (verticale) et par $\{b_j/1 \leqslant j \leqslant J\}$ la marge ligne (horizontale). On suppose -sans aucunement restreindre la généralité- que, pour tout i de I (resp. j de J) $a_i \neq 0$ (resp. $b_j \neq 0$). On a :

$$a_i = \sum_{1 \leqslant j \leqslant J} c_{ij},\quad b_j = \sum_{1 \leqslant i \leqslant I} c_{ij},\ \text{pour tout } (i,j) \text{ de } I \times J, \text{ d'autre part},$$

$$n = \sum_{1 \leqslant i \leqslant I} a_i = \sum_{1 \leqslant j \leqslant J} b_j .$$

En d'autres termes $\{c_{ij}/1 \leqslant i \leqslant I,\ 1 \leqslant j \leqslant J\}$ définit un croisement entre les deux partages ordonnés $\{a_i/1 \leqslant i \leqslant I\}$ et $\{b_j/1 \leqslant j \leqslant J\}$ de l'entier n.

Le problème consiste à construire <u>la valeur maximale exacte</u> de la somme des carrés des nombres entiers contenus dans les cases, <u>sous contrainte des marges fixées</u> ; soit

$$\text{Max} \sum \{c_{ij}^2/1 \leqslant i \leqslant I,\ 1 \leqslant j \leqslant J\} \qquad (1)$$

sous les contraintes :

$$\sum_{1\leqslant j\leqslant J} c_{ij}=a_i,\ 1\leqslant i\leqslant I,$$

$$\sum_{1\leqslant i\leqslant I} c_{ij}=b_j,\ 1\leqslant j\leqslant J. \qquad (2)$$

où les a_i et les b_j sont fixés.

Ce problème est reconnu comme "très difficile" voire même "quasi impossible" à résoudre. Citons en effet à cet égard L. Hubert et P. Arabie |HUBERT & ARABIE(1985)| d'une part et F. Marcotorchino |MARCOTORCHINO(1984)| d'autre part :

"...Constructing an exact bound, conditional on the fixed row and column totals of the given contingency table, is a very difficult problem of combinatorial optimization..."

..."la maximisation...relève de l'utilisation d'algorithmes de programmation non-linéaire et ce, si l'on ne suppose de conditions d'intégrité des n_{uv} (c_{ij} dans nos notations) (<u>cas où ce problème serait quasiment impossible à résoudre</u>)"...

C'est précisément ce problème que nous résolvons ici, dans toute sa généralité.

Commençons par exprimer de façon très intuitive les quelques idées germes qui ont permis de faire pousser cette recherche.

La première idée consiste à remplacer la notion de formule mathématique -dont l'argument serait le couple de partages $[\{a_i/1\leqslant i\leqslant I\},\ \{b_j/1\leqslant j\leqslant J\}]$- par celle beaucoup plus générale, d'algorithme.

Ce dernier sera défini au niveau le plus synthétique du tableau de contingence et aura un caractère récursif.

Le point de départ est le tableau vide à l'intérieur, mais ayant ses marges remplies qu'il s'agit de répartir de façon compatible et optimale. Il serait étonnant que dans cette structure optimale, il n'y ait pas un couple d'indices (i_o,j_o) tel que

$c_{i_oj_o}=\min(a_{i_o},b_{j_o})$;

ce qui correspond au "déchargement" de a_{i_o} dans la colonne j_o si $a_{i_o}\leq b_{j_o}$ (resp. de b_{j_o} dans la ligne i_o si $a_{i_o}\geq b_{j_o}$). On s'attend d'ailleurs à ce que $c_{i_oj_o}$ soit relativement "gros" et $|a_{i_o}-b_{j_o}|$ relativement "petit".

Si nous arrivons à caractériser la définition mathématique d'un tel couple (i_o,j_o) et à le localiser dans un ensemble de faible taille de $\{1,2,\ldots,I\}x\{1,2,\ldots,J\}$, nous aurons gagné. En effet, le chargement de la case (i_o,j_o) permet la suppression de la ligne i_o si $a_{i_o}\leq b_{j_o}$ (resp. de la colonne j_o si $b_{j_o}\leq a_{i_o}$). La dimension du problème se trouve ainsi diminuée de IxJ à (I-1)xJ [resp. Ix(J-1)] ou même (I-1)x(J-1) dans le cas où $a_{i_o}=b_{j_o}$. En effet, la suppression d'une ligne (resp. colonne) et le réaménagement conséquent de la marge colonne (resp. ligne) d'une configuration optimale, donne une configuration optimale.

- Nous n'en dirons pas plus au niveau de cette introduction pour ce qui concerne l'approche algorithmique qui fournit une solution au problème posé.

A défaut de connaître la valeur exacte de (1) différents auteurs ont proposé des bornes fonctions symétriques de $(\{a_i/1\leq i\leq I\},\{b_j/1\leq j\leq J\})$. Nous montrons que les plus intéressantes résultent de l'application de l'inégalité de Shwartz, mais conçue dans un cadre logique. Nous déterminons alors l'inégalité de Shwartz qui donne la majoration la plus serrée possible, laquelle est sensiblement meilleure que celles fournies dans [HUBERT & ARABIE (1985)].

De toute façon, nous démontrons que la borne, dont l'expression est dissymétrique définie par

$$\min\Big(\sum_{1\leq i\leq I} a_i^2,\ \sum_{1\leq j\leq J} b_j^2\Big)$$

est elle-même sensiblement meilleure que celle définie par la "meilleure" inégalité de Shwartz. Ce résultat est en quelque sorte intermédiaire entre l'approche "formule mathématique" que nous analyserons en premier et celle -la plus importante- algorithmique dont nous apporterons la meilleure démonstration mathématique dans la deuxième partie de cette étude.

Il s'agit maintenant de répondre à la question : quel est l'intérêt de trouver le maximum de

$$\sum\{c^2_{ij}/(i,j) \in I\times J\} \qquad (3)$$

-où $I=\{1,2,\ldots,I\}$ et $J=\{1,2,\ldots,J\}$- ; davantage et plus précisément, quel est l'intérêt de la découverte d'une structure maximale d'association entre les marges ligne et colonne, au sens du critère (3).

Beaucoup d'indices d'association entre deux variables-partition (qualitatives nominales) sont conçus à partir de (3). Même si ces derniers supposent une certaine normalisation de type formel ou statistique mettant en oeuvre et de façon symétrique $\sum_i a^2_i$ et $\sum_j b^2_j$, il importe de pouvoir reconnaître, pour corriger la normalisation, la valeur maximale pouvant être atteinte par ces indices. Cette dernière dépend directement de la valeur maximale de (3).

Maintenant, la découverte d'une structure maximale d'association peut être d'un grand intérêt dans tout problème d'"estimation" de l'évolution d'une classification, à condition d'admettre un principe de parcimonie, relativement au critère (3). Plus précisément, connaissant les types $(a_1,\ldots,a_i,\ldots,a_I)$ et $(b_1,\ldots,b_j,\ldots,b_J)$ de deux partitions différentes d'un ensemble O d'objets ou individus, quelle est la table de contingence permettant de passer de l'une des répartitions définie par le premier type, à l'autre définie par le deuxième type, de façon à maximiser le critère (3).

On rencontre ce problème dans toute étude de mobilité sociale. Ainsi, pour un corps électoral donné et relativement à deux élections législatives consécutives, on connaît simplement -pour chacune des élections- les nombres de voix respectivement concédés aux différents partis en présence. Le problème est alors d'imaginer la relation de passage, matérialisée par le tableau de contingence dont les marges définissent respectivement les résultats des deux élections. Mais peut-être que pour ce dernier problème, le seul axiome de parcimonie dans le passage d'une structure partition à une autre, ne suffit pas. S'il suffit de plus de tenir compte d'une échelle sous-jacente, c'est-à-dire, d'un ordre total sur l'ensemble des lignes (resp. colonnes), c'est un problème auquel nous proposons une solution [LERMAN (sous presse)].

II. BORNES OBTENUES PAR UNE FORMULE

II.1. Introduction ; Famille de bornes résultant de l'inégalité de Shwartz.

Nous reprenons nos notations du paragraphe I et -pour matérialiser les calculs- nous considérons la table $T=\{c_{ij}/1\leq i\leq I,\ 1\leq j\leq J\}$ de contingence comme celle des cardinaux des classes du croisement de deux partitions en classes étiquetées P et Q de types respectifs $(a_1,a_2,\ldots,a_I)$ et $(b_1,b_2,\ldots,b_J)$, sur un ensemble E d'objets. c_{ij} est le cardinal de la classe E_{ij} résultant de l'intersection de la classe $E_{i.}$ de la partition P et de la classe $E_{.j}$ de la partition Q, $1\leq i\leq I$, $1\leq j\leq J$.

Nous désignerons par $F=P_2(E)$ l'ensemble des paires d'objets de E et par G=ExE l'ensemble des couples d'objets de E. RF(P) [resp.RF(Q)] le sous-ensemble des paires d'objets réunies par la partition P(resp.Q). De même RG(P) [resp.RG(Q)] est le sous-ensemble des couples d'objets réunis par la partition P (resp.Q). On introduit enfin $RF(P\wedge Q)$ et $RG(P\wedge Q)$ qui sont respectivement définis par le sous-ensemble des paires d'objets et par celui des couples d'objets réunis par la partition croisée $P\wedge Q$. On a

$$RF(P\wedge Q)=RF(P)\cap RF(Q)$$

et

$$RG(P\wedge Q)=RG(P)\cap RG(Q).$$

Introduisons à présent φ(resp.ψ) la fonction indicatrice de RF(P) [resp. RF(Q)]. De même, soit ρ(resp.τ) la fonction indicatrice de RG(P) [resp.RG(Q)]. Dans ces conditions, la fonction indicatrice de $RF(P\wedge Q)$ est $\varphi.\psi$ et celle de $RG(P\wedge Q)$ est $\rho.\tau$.

En ce qui concerne les cardinaux, on a

$$\mathrm{card}[RF(P)]=\sum\{\binom{a_i}{2}/1\leq i\leq I\}$$

$$=\frac{1}{2}\left(\sum_i a_i^2-n\right) \qquad (1)$$

$$\mathrm{card}\left[RF(Q)\right] = \sum\left\{\binom{b_j}{2} / 1 \leqslant j \leqslant J\right\}$$

$$= \frac{1}{2}\left(\sum_j b_j^2 - n \right) \quad ,$$

$$\mathrm{card}\left[RF(P\,Q)\right] = \sum\left\{\binom{c_{ij}}{2} / 1 \leqslant i \leqslant I,\ 1 \leqslant j \leqslant J\right\}$$

$$= \frac{1}{2}\left(\sum_{(i,j)} c_{ij}^2 - n\right). \quad (2)$$

D'autre part,

$$\mathrm{card}\left[RG(P)\right] = \sum_i a_i^2$$

$$\mathrm{card}\left[RG(Q)\right] = \sum_j b_j^2 \quad , \qquad (3)$$

$$\mathrm{card}\left[RG(P \wedge Q)\right] = \sum_{(i,j)} c_{ij}^2 \ . \qquad (4)$$

On a donc

$$\mathrm{card}\left[RG(.)\right] = 2\mathrm{card}\left[RF(.)\right] + n. \qquad (5)$$

Il est ainsi clair que la majoration de $s(T) = \sum\{c_{ij}^2 / 1 \leqslant i \leqslant I,\ 1 \leqslant j \leqslant J\}$ est équivalente à celle de

$$r(T) = \sum\left\{\binom{c_{ij}}{2} / 1 \leqslant i \leqslant I,\ 1 \leqslant j \leqslant J\right\}.$$

Pour cette dernière, L. Hubert et Ph. Arabie [HUBERT & ARABIE(1985)] proposent deux bornes ayant un caractère symétrique par rapport aux marges.

La première est définie par

$$\frac{1}{2}\left[\left\{\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}\right)\right] \qquad (6)$$

et la seconde par

$$\sqrt{\left[\sum_i \binom{a_i}{2}\right]\left[\sum_j \binom{b_j}{2}\right]} \ . \qquad (7)$$

La première correspond à la moyenne arithmétique et la seconde, à celle géométrique de card[RF(P)] et de card[RF(Q)]. Mais comme la moyenne géométrique de deux nombres positifs est inférieure à celle, arithmétique, la borne (7) est plus proche de la borne exacte que celle (6).

En utilisant les fonctions indicatrices φ et ψ, l'inégalité peut se mettre sous la forme

$$\sum_{p\in F} \varphi(p)\psi(p) \leqslant \sqrt{\left[\sum_{p} \varphi(p)\right]\left[\sum_{p} \psi(p)\right]}. \quad \textbf{(8)}$$

Il s'agit d'une inégalité de type Shwartz qui est obtenue en écrivant que le discriminant du trinôme suivant du second degré en x, est négatif :

$$\sum_{p\in F} \left[x\varphi(p)-\psi(p)\right]^2. \quad \textbf{(9)}$$

Mais il existe toute une famille d'inégalités de même type obtenue par la considération de la famille des trinômes du second degré en x :

$$\sum_{p\in F} \{x\left[\varphi(p)-\mu\right]-\left[\psi(p)-\nu\right]\}^2 . \quad \textbf{(10)}$$

Dans ces conditions, la question se pose de savoir quelles sont les valeurs de μ et de ν (paramètres réels) pour lesquelles la borne est la plus serrée.

D'autre part, on peut obtenir des inégalités de même type -cette fois-ci directement pour s(T)- en travaillant au niveau de l'ensemble G des couples (au lieu de l'ensemble F des paires on écrira alors qu'est négatif le discriminant du trinôme du second degré en x suivant

$$\sum_{q\in G} \{x\left[\rho(q)-\mu\right]-\left[\tau(q)-\nu\right]\}^2 . \quad \textbf{(11)}$$

Nous commencerons par montrer que c'est à partir de **(10)** -c'est-à-dire, en travaillant au niveau de F- qu'on obtient pour μ et ν fixés, la borne la plus petite.

Dans un deuxième temps, nous déterminerons, dans le cadre de la famille des inégalités obtenues à partir de **(10)** et indexée par (μ,ν), les valeurs précises de μ et de ν qui donnent la meilleure borne au moyen d'une formule symétrique où -de façon basique- seules interviennent l'addition et la multiplication. Nous appellerons "analytique" une telle borne.

II.2. La meilleure borne "analytique"

A partir de (**10**) et de (**11**), on obtient (en écrivant que le discriminant du trinôme du second degré en x est négatif :

$$\sum_{p\in F} [\varphi(p)-\mu][\psi(p)-\nu]$$

$$\leq \{(\sum_{p\in F} [\varphi(p)-\mu]^2)(\sum_{p\in F} [\psi(p)-\nu]^2)\}^{\frac{1}{2}} \qquad \textbf{(12)}$$

et

$$\sum_{q\in G} [\rho(q)-\mu][\tau(q)-\nu]$$

$$\leq \{(\sum_{q\in G} [\rho(q)-\mu]^2)(\sum_{q\in G} [\tau(q)-\nu]^2)\}^{\frac{1}{2}}. \qquad \textbf{(13)}$$

La relation (**12**) conduit à

$$s(T)\leq \mu\sum_j b_j^2+\nu\sum_i a_i^2+n(1-\mu-\nu)$$

$$+\{[n^2\mu^2+(-2\mu+1)\sum_i a^2_i-n(1-\mu)^2]$$

$$\times[n^2\nu^2+(-2\nu+1)\sum_j b_j^2-n(1-\nu)^2]\}^{\frac{1}{2}}$$

En désignant par A, $\sum_i a_i^2$ (resp.B, $\sum_j b_j^2$), on écrit :

$$s(T)\leq \mu B+\nu A-n^2\mu\nu+n(1-\mu)(1-\nu)$$

$$+\{[n^2\mu^2+(-2\mu+1)A-n(1-\mu)^2]$$

$$\times[n^2\nu^2+(-2\nu+1)B-n(1-\nu)^2]\}^{\frac{1}{2}}. \qquad \textbf{(14)}$$

La relation (**13**) conduit à

$$s(T)\leq \mu B+\nu A-n^2\mu\nu$$

$$+\{[n^2\mu^2+(-2\mu+1)A][n^2\nu^2+(-2\nu+1)B]\}^{\frac{1}{2}}. \qquad \textbf{(15)}$$

Comparer les bornes fournies par (**14**) et (**15**) revient à comparer

$$\{[\Lambda-n(1-\mu)^2][M-n(1-\nu)^2]\}^{\frac{1}{2}} + n(1-\mu-\nu+\mu\nu)$$

avec $\sqrt{\Lambda M}$ où nous avons posé :

$$\Lambda = n^2\mu^2+(-2\mu+1)A$$

et

$$M = n^2\nu^2+(-2\nu+1)B.$$

Comparer les bornes fournies par **(15)** et **(14)** revient à comparer respectivement $(\sqrt{\Lambda M}-n\lambda)$ et $\left[\Lambda-n(1-\mu)^2\right]\left[M-n(1-\nu)^2\right]$, où on a au préalable posé $\lambda=(1-\mu)(1-\nu)$. $\Lambda \geq n(1-\mu)^2$ et $M \geq n(1-\nu)^2$ conduisent à $\sqrt{\Lambda M} \geq n\lambda$. De sorte qu'il y a lieu de comparer $(\sqrt{\Lambda M}-n\lambda)^2$ avec $\left[\Lambda-n(1-\mu)^2\right]\left[M-n(1-\nu)^2\right]$.

La deuxième expression se met sous la forme :
$\Lambda M+n^2\lambda^2-nM(1-\mu)^2-n\Lambda(1-\nu)^2$.

Il reste alors à comparer $-2n\lambda\sqrt{\Lambda M}$ avec $-nM(1-\mu)^2-n\Lambda(1-\nu)^2$. Or $nM(1-\mu)^2+n\Lambda(1-\nu)^2-2n\lambda\sqrt{\Lambda M} = \left[\sqrt{nM}(1-\mu)-\sqrt{n\Lambda}(1-\nu)\right]^2 \geq 0$,

Ainsi, la borne définie par le second membre de **(14)** est plus serrée que celle définie par le second membre de **(15)**. D'où, la propriété :

PROPRIETE

Si T est regardé comme le tableau de contingence du croisement de deux partitions sur un ensemble E d'objets, l'inégalité de Shwartz conçue au niveau de l'ensemble des paires d'objets de E (i.e. des parties à deux éléments de E), donne une borne plus précise que celle conçue au niveau de l'ensemble ExE des couples d'objets de E.

Nous allons maintenant reprendre le second membre de **(14)** et nous allons chercher à déterminer les valeurs de μ et de ν pour lesquelles il est le plus petit. Si nous désignons par $\phi(\mu,\nu)$ ce second membre de **(14)** moins n et divisé par $m=n(n-1)$, on peut -après mise en forme- écrire :

$$\Phi(\mu,\nu) = \mu\beta+\nu\alpha-\mu\nu+\sqrt{UV} \qquad \textbf{(16)}$$

où

$$\alpha = \frac{(A-n)}{n(n-1)} = \frac{\left[\sum_i a_i(a_i-1)\right]}{n(n-1)} = \frac{\mathrm{card}\left[RF(P)\right]}{\mathrm{card}(F)} ,$$

$$\beta = \frac{(B-n)}{n(n-1)} = \frac{\left[\sum_j b_j(b_j-1)\right]}{n(n-1)} = \frac{\mathrm{card}\left[RF(Q)\right]}{\mathrm{card}(F)},$$

$$U = \mu^2 - (2\mu-1)\alpha \quad \text{et} \quad V = \nu^2-(2\nu-1)\beta.$$

En écrivant que la dérivée par rapport à μ est nulle, on obtient la relation suivante :

$$(\beta-\nu)-\sqrt{\frac{V}{U}}(\alpha-\mu) = 0. \quad \textbf{(17)}$$

Une solution apparaît clairement, elle est définie par :

$$\mu = \alpha \text{ et } \nu = \beta. \quad \textbf{(18)}$$

On voit que l'annulation de la dérivée par rapport à ν de la fonction $\Phi(\mu,\nu)$ conduit à la même solution pour μ et ν, laquelle définit un extrémum local de $\Phi(\mu,\nu)$. Nous allons précisément montrer qu'il s'agit d'un minimum général. On a

$$\Phi(\alpha,\beta)=(\alpha\beta+\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}), \quad \textbf{(19)}$$

où nous avons noté $\bar{\alpha}$(resp.$\bar{\beta}$) $(1-\alpha)$ $\left[\text{resp.}(1-\beta)\right]$.

Nous avons ainsi à établir que

$$\Phi(\mu,\nu) \geq \Phi(\alpha,\beta). \quad \textbf{(20)}$$

Au préalable, mettons U et V sous la forme suivante :

$$U=(\mu-\alpha)^2+\alpha\bar{\alpha} \text{ et } V=(\nu-\beta)^2+\beta\bar{\beta}.$$

On a ainsi à établir que

$$\mu\beta+\nu\alpha-\mu\nu+\{\left[(\mu-\alpha)^2+\alpha\bar{\alpha}\right]\left[(\nu-\beta)^2+\beta\bar{\beta}\right]\}^{\frac{1}{2}}$$
$$\geq \alpha\beta+\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}. \quad \textbf{(21)}$$

Cela revient à dire que

$$(\mu-\alpha)(\nu-\beta)+\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}} \leq$$
$$\{\left[(\mu-\alpha)^2+\alpha\bar{\alpha}\right]\left[(\nu-\beta)^2+\beta\bar{\beta}\right]\}^{\frac{1}{2}}. \quad \textbf{(22)}$$

Donc que

$$(\mu-\alpha)^2(\nu-\beta)^2+\alpha\bar{\alpha}\beta\bar{\beta}+2\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}\,(\mu-\alpha)(\nu-\beta)$$

$$\leqslant[(\mu-\alpha)^2+\alpha\bar{\alpha}]\,[(\nu-\beta)^2+\beta\bar{\beta}]. \qquad (23)$$

Cette inégalité se ramène à

$$2\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}\,(\mu-\alpha)(\nu-\beta)\leqslant(\mu-\alpha)^2\beta\bar{\beta}+(\nu-\beta)^2\alpha\bar{\alpha}, \qquad (24)$$

laquelle étant aquise. D'où le théorème :

<u>THEOREME 1</u>. La meilleure borne analytique pour $s(T)=\sum\{c_{ij}^2/(i,j)\}$, découlant de l'inégalité de Shwartz, est donnée par

$$n(n-1)\left[\alpha\beta+\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}\right]+n, \qquad (25)$$

où $\alpha=\sum_i a_i(a_i-1)/n(n-1)$ $\left[\text{resp. } \beta=\sum_j b_j(b_j-1)/n(n-1)\right]$ et où $\bar{\alpha}=1-\alpha$ (resp. $\bar{\beta}=1-\beta$).

II.3. Situation de la borne définie par min(A,B)

Nous nous proposons ici de situer la borne évidente mais dissymétrique définie par

$$\min\left(\sum_i a_i^2, \sum_j b_j^2\right), \qquad (26)$$

par rapport à celle "analytique" définie par (25). Il revient au même de comparer $\min(\alpha,\beta)$ avec $\Phi(\alpha,\beta)=\left[\alpha\beta+\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}\right]$.

On ne restreint en rien la généralité si on suppose $\alpha\leqslant\beta$. Cette dernière inégalité est équivalente à celle

$$\alpha(1-\beta)\leqslant(1-\alpha)\beta\ , \qquad (27)$$

qui se met sous la forme

$$\alpha\bar{\beta}\leqslant\alpha\bar{\beta}$$

et qui implique

$$\sqrt{\alpha\bar{\beta}}\leqslant\sqrt{\bar{\alpha}\beta}\ . \qquad (28)$$

En multipliant les deux membres par $\sqrt{\alpha\bar{\beta}}$, on obtient

$$\alpha\bar{\beta} \leqslant \sqrt{\alpha\bar{\alpha}\beta\bar{\beta}}$$

En ajoutant maintenant $\alpha\beta$ aux deux membres, on obtient

$$\alpha\beta+\sqrt{\alpha\bar{\alpha}\beta\bar{\beta}} \geqslant \alpha(\bar{\beta}+\beta)=\alpha,$$

soit

$$\Phi(\alpha,\beta)\geqslant\alpha.$$

D'où le théorème :

THEOREME 2. Relativement au problème de la majoration de $s(T)=\sum\{c_{ij}^2/(i,j)\}$, la borne définie par $\min(\sum_i a_i^2,\sum_j b_j^2)$ est inférieure [serre de plus près $s(T)$] que la plus petite borne analytique résultant de l'application de l'inégalité de Shwartz.

DEFINITION. En désignant par $\mathbb{I}$ l'ensemble des indices $\{1,2,\ldots,i,\ldots,I\}$, le partage de n défini par $(a_1,a_2,\ldots,a_i,\ldots,a_I)$ est dit "plus fin" que celui défini par $(b_1,b_2,\ldots,b_j,\ldots,b_J)$, s'il existe une partition $\{\mathbb{I}_j/1\leqslant j\leqslant J\}$ de $\mathbb{I}$ en J classes non vides, telle que :

$$(\forall j,1\leqslant j\leqslant J),\ b_j=\sum\{a_i/i\in\mathbb{I}_j\}. \quad (30)$$

PROPRIETE 3. Si parmi les deux partages de n $\{a_i/1\leqslant i\leqslant I\}$ et $\{b_j/1\leqslant j\leqslant J\}$, l'un est plus fin que l'autre, la borne définie par $\min(\sum_i a_i^2,\sum_j b_j^2)$ est exacte.

Ce résultat est trivial ; mais nous avons voulu le ponctuer car il représente une transition entre l'approche que nous avons appelée "analytique" et qui résulte de l'application de l'inégalité de Shwartz et celle "algorithmique" que nous allons entreprendre au paragraphe suivant.

Etablissons plus explicitement ce résultat. On supposera pour cela que c'est le partage $\{a_i/1\leqslant i\leqslant I\}$ qui est plus fin que celui $\{b_j/1\leqslant j\leqslant J\}$. Il est clair que

$$\sum_{(i,j)} c_{ij}^2 \leqslant \sum_i a_i^2. \quad (31)$$

D'autre part, on répartira la marge colonne b_j sur $\mathbb{I}_j$, conformément à la décomposition (30) ci-dessus. De sorte que $c_{ij}=a_i$ si $i\in\mathbb{I}_j$. Par conséquent, l'inégalité (31) se réduit à une égalité.

III. JUSTIFICATION DE L'APPROCHE ALGORITHMIQUE

III.1. Définitions, outils, exemples et contre-exemples.

On commencera par reprendre ici le début de l'introduction (§I) qui comprend -encore de façon très vague- le germe de l'idée algorithmique. Commençons par préciser les notations. Sans aucunement restreindre la généralité, on supposera que

$$a_1 \geqslant a_2 \geqslant \ldots \geqslant a_i \geqslant \ldots \geqslant a_I$$

$$b_1 \geqslant b_2 \geqslant \ldots \geqslant b_j \geqslant \ldots \geqslant b_J \qquad (1)$$

et que

$$a_1 \geqslant b_1 .$$

De sorte qu'on désignera par $\mathcal{C}$ l'ensemble des tables de contingence T pour lesquelles on a (1). D'autre part, on dira que T est de type $(a_1,a_2,\ldots,a_i,\ldots,a_I\ ;\ b_1,b_2,\ldots,b_j,\ldots,b_J)$. (2)

Comme au paragraphe précédent, relativement à un tableau $T=\{c_{ij}/1\leqslant i\leqslant I,\ 1\leqslant j\leqslant J\}$, on désignera par $s(T)$ la somme des carrés :

$$s(T)=\sum\{c_{ij}^2/1\leqslant i\leqslant I,\ 1\leqslant j\leqslant J\}. \quad (3)$$

On désignera d'autre part par

$$s_m\begin{bmatrix} a_1,\ldots,a_i,\ldots,a_I \\ \\ b_1,\ldots,b_j,\ldots,b_J \end{bmatrix} \qquad (4)$$

la valeur maximale pouvant être atteinte de (3) pour un tableau de type (2).

Comme nous l'avons déjà signalé dans l'introduction, pour aboutir à la structure T_m de la table de contingence qui donne lieu à la valeur (4), on partira du tableau vide à l'intérieur, mais ayant ses marges remplies qu'il s'agit de répartir de façon compatible et optimale. Après l'avoir justifié, on sera conduit -à chaque pas à installer le contenu d'une marge ligne α_i dans une colonne j (auquel cas $\alpha_i \leq \beta_j$: contenu de la marge colonne j), ou bien le contenu d'une marge colonne β_j dans une ligne i (auquel cas $\alpha_i \leq \beta_j$). Au départ $\alpha_i = a_i$ pour $1 \leq i \leq I$ (resp. $\beta_j = b_j$ pour $1 \leq j \leq J$). On dira qu'on "vide" ("déverse" ou "décharge") α_i dans la colonne j, ou bien β_j dans la ligne i.

Dans l'ensemble des couples LignesxColonnes, tout le problème est de reconnaître à chaque fois le couple (ligne, colonne) pour lequel il y a lieu de réaliser le déchargement. Il y a tout au moins lieu de reconnaître un sous-ensemble de faible cardinal de l'ensemble des couples, où doit se réaliser le déchargement.

La "résolution" d'un premier couple (i_o,j_o) -qui correspond au déversement de a_{i_o} dans j_o si $a_{i_o} \leq b_{j_o}$ (resp. de b_{j_o} dans i. si $a_{i_o} > b_{j_o}$)- doit intuitivement parlant, être telle que $c_{i_oj_o} = \min(a_{i_o}, b_{j_o})$ relativement "gros" et $|a_{i_o} - b_{j_o}|$ relativement "petit".

Comme nous l'avons mentionné dans l'introduction, la résolution d'un couple diminue la dimension du problème en diminuant d'une unité, voire même de deux (si les deux arguments du couple résolu sont identiques) le cardinal suivant :

(nombre de lignes + nombre de colonnes).

Nous avons pu croire que le premier couple (a_{i_o}, b_{j_o}) à résoudre était celui qui réalisait la valeur absolue de "la plus petite différence" ($|a_{i_o} - b_{j_o}|$ minimal) et en cas d'exaequos sur ce dernier critère, un de ceux qui réalisaient la plus grande somme $[(a_{i_o} + b_{j_o})$ maximal$]$. Le contre exemple suivant nous montre qu'il n'en est rien.

$T_1 =$

1	3	3	3	10
7	0	0	0	7
8	3	3	3	17

$s(T_1) = 77$

$T_2=$

8	2	0	0	10
0	1	3	3	7
8	3	3	3	17

$s(T_2)=87$,

où la deuxième configuration correspond à l'optimum.

S'il faut suivre l'exemple précédent et croire que le premier couple (a_{i_o}, b_{j_o}) à résoudre est celui qui réalise "la plus grande somme" $[(a_{i_o}+b_{j_o})$ maximal$]$ et en cas d'exaequos, un de ceux qui réalisent la valeur minimale de la valeur absolue de la différence $[a_{i_o}-b_{j_o}]$, le contre exemple suivant nous détrompe :

$T_3=$

8	0	2	10
0	5	2	7
0	0	1	1
8	5	5	18

$s(T_3)=98$

$T_4=$

0	5	5	10
7	0	0	7
1	0	0	1
8	5	5	

$s(T_4)=100$,

où la configuration T_4 pour laquelle il faut commencer par "la plus petite différence" est optimale.

On peut croire alors qu'il suffit d'essayer deux chaînes de traitement où dans le premier cas, on procède avec le critère de "la plus petite différence" et où, dans le second cas on procède avec le critère de "la plus grande somme". Mais alors, si on considère le type suivant d'un tableau T :

(100,70,10,10,7,1 ; 80,50,50,8,3,3,2,2),

l'application de la première stratégie conduit à la configuration suivante du tableau :

T_5=

0	50	50	0	0	0	0	0	100
70	0	0	0	0	0	0	0	70
10	0	0	0	0	0	0	0	10
0	0	0	0	3	3	2	2	10
0	0	0	7	0	0	0	0	7
0	0	0	1	0	0	0	0	1
80	50	50	8	3	3	2	2	198

$s(T_5)=10076$

L'application de la deuxième stratégie conduit à la configuration suivante du tableau :

T_6=

80	0	20	0	0	0	0	0	100
0	50	20	0	0	0	0	0	70
0	0	10	0	0	0	0	0	10
0	0	0	8	0	0	2	0	10
0	0	0	0	3	3	0	1	7
0	0	0	0	0	0	0	1	1
80	50	50	8	3	3	2	2	198

$s(T_6)=9888$

En réalité, la configuration optimale correspond au tableau T_7 suivant qui résulte de l'application de la première stratégie au sous tableau de type (100,70,10 ; 80,50,50) et de celle de la seconde stratégie au sous tableau de type (10,7,1;8,3,3,2,2). En effet, le type de notre tableau se décompose en le produit des deux derniers types.

T_7=

0	50	50	0	0	0	0	0	100
70	0	0	0	0	0	0	0	70
10	0	0	0	0	0	0	0	10
0	0	0	8	0	0	0	0	10
0	0	0	0	3	3	2	1	7
0	0	0	0	0	0	0	1	1
80	50	50	8	3	3	2	2	198

$s(T_7)=10088$.

D'où l'idée de définir sur l'ensemble des couples :
$\{(a_i,b_j)/1\leqslant i\leqslant I, 1\leqslant j\leqslant J\}$ **(5)**

une relation de préordre partiel résultant de l'intersection de deux préordres totaux où le premier ω_d est conforme à la différence et où le second ω_s est conforme à la somme :

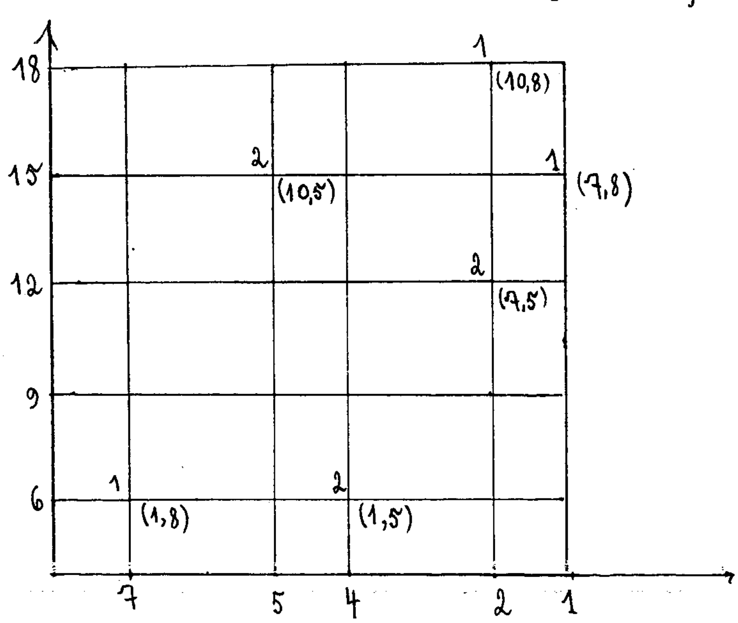
$$(\forall[(i,j),(i',j')]),(a_{i'},b_{j'})\leqslant(a_i,b_j) \text{ (pour } \omega_d)$$

$$\Leftrightarrow |a_i-b_j|\leqslant|a_{i'}-b_{j'}| \quad \textbf{(6)}$$

$$(\forall[(i,j),(i',j')]),(a_{i'},b_{j'})\leqslant(a_i,b_j) \text{ (pour } \omega_s)$$

$$\Leftrightarrow (a_{i'}+b_{j'})\leqslant(a_i+b_j). \quad \textbf{(7)}$$

Ainsi, pour le type suivant d'un tableau
(10,7,1 ; 8,5,5),
rapportons N^2 à deux axes où l'axe horizontal (resp. vertical) est relatif à l'échelle de la différence (resp. somme) entre un a_i et un b_j ($1\leqslant i\leqslant I$, $1\leqslant j\leqslant J$).



Sur ce diagramme nous avons marqué les différents couples (a_i,b_j) ainsi que leurs fréquences respectives (en haut et à gauche des points qui les représentent). Il y a -pour le préordre intersection des deux préordres totaux (ω_d et ω_s)- deux couples extrémaux : (7,8) et (10,8).

L'algorithme que nous allons proposer repose sur la simple propriété suivante :

"<u>La solution optimale</u> peut être obtenue en commençant par la résolution d'un couple extrémal (a_i,b_j) [déversement de a_i dans la colonne j (resp. de b_j dans la ligne i) si $a_i \leq b_j$ (resp. $b_j < a_j$)]".

De toute façon, nous démontrons (cf.§III.3.) que le plus grand entier du tableau définissant la configuration optimale, correspond bien à la résolution d'un couple. Le contraire de la propriété suppose que dans cette configuration optimale, aucun des couples résolus n'est extrémal ! Cela, on en convient semble impossible.

Toutefois, en toute rigueur et dans le cas le plus général (plusieurs couples extrémaux), cette propriété reste encore (au niveau de ce texte) à l'état de conjecture.

Mais, dans le cas où <u>une même part marginale se retrouve</u> des deux côtés du tableau [$a_i=b_j$, pour un (i,j)], la solution optimale passe nécessairement par la résolution de ce couple (a_i,b_j) (cf. §III.2.).

D'autre part, c'est tout à fait complètement que nous établissons que dans le cas où il y a <u>un seul couple extrémal,</u> la configuration optimale passe nécessairement par la résolution de ce couple. Il s'agit certainement du résultat qui nous a demandé le plus d'efforts pour être démontré (cf. §III.3.).

Enfin, dans <u>quelques situations particulières où il y a plus d'un seul couple extrémal,</u> nous montrons que la solution optimale passe par la résolution d'un couple extrémal (cf. §III.4.).

Deux types de transformation d'un tableau $T=\{c_{ij}/(i,j) \in I \times J\}$ sont à considérer comme outils pour la preuve.

La première correspond à une notion d'"opération élémentaire" préservant le type du tableau T, dépendant d'un quadruplet d'indices (i,i',j,j') et d'un entier k inférieur à min $(c_{ij},c_{i'j'})$. Cette transformation que nous notons $\varphi^{k}_{(i,i',j,j')}$ a pour effet de remplacer $(c_{ij},c_{ij'},c_{i'j},c_{i'j'})$ repectivement par $(c_{ij}-k,c_{ij'}+k,c_{i'j}+k,c_{i'j'}-k)$, selon le schéma suivant :

$$\begin{tikzcd} & j & j' \\ i & (c_{ij}-k) \arrow[r] \arrow[d] & (c_{ij'}+k) \\ i' & (c_{i'j}+k) & (c_{i'j'}-k) \arrow[l] \arrow[u] \end{tikzcd}$$

On a

$$s\left[\varphi^{k}_{(i,i',j,j')}(T)\right]=s(T)+2k\left[2k+(c_{ij'}+c_{i'j}-c_{ij}-c_{i'j'})\right] \qquad (8)$$

La deuxième transformation que nous notons ψ^{k}_{ij} change le type du tableau T et n'est possible que si $c_{ij}\geqslant k$. Elle consiste tout simplement à enlever k au contenu c_{ij} de la case (i,j) du tableau. $\psi^{k}_{ij}(T)$ est un tableau de type $(a_1,a_2,\ldots,a_{(i-1)},(a_i-k),a_{i+1},\ldots,a_I\ ;\ b_1,b_2,\ldots,b_{(j-1)},(b_j-k),b_{j+1},\ldots,b_J)$, où la position de (a_i-k) $\left[\text{resp.}(b_j-k)\right]$ a à être reconsidérée par rapport aux $a_{i'}$ (resp. $b_{j'}$).

On a

$$s\left[\psi^{k}_{ij}(T)\right]=s(T)-k(2c_{ij}-k). \qquad (9)$$

Donnons à présent l'idée générale de l'algorithme qui sera techniquement précisée au paragraphe IV. Nous terminerons ce paragraphe en en donnant un exemple d'application.

Partant du tableau T_o à I lignes et J colonnes, dont seules les marges se trouvent remplies, l'idée générale consiste à explorer tous les chemins possibles dont chaque pas consiste en la résolution d'un couple extrémal d'un tableau recueilli après suppression d'une ligne ou d'une colonne, résultant de la résolution d'un couple extrémal au pas précédent.

Ainsi, la première étape de l'algorithme consiste à déterminer l'ensemble des couples extrémaux de T_o.

La deuxième étape consiste à ranger les éléments de cet ensemble conformément à ω_d (par valeurs croissantes du critère de la différence).

Un résultat très important pour la marche de l'algorithme consiste à montrer que si (a_{i1},b_{j1}) et (a_{i2},b_{j2}) sont deux couples extrémaux tels que

$|a_{i1}-b_{j1}|<|a_{i2}-b_{j2}|$ $\left[\text{et donc } (a_{i2}+b_{j2})>(a_{i1}+b_{j1})\right]$, la résolution du couple extrémal à une certaine étape préserve le caractère extrémal de (a_{i2},b_{j2}) pour le tableau qui en résulte (cf.§ IV).

Il suffit donc d'empiler la suite ordonnée des couples (conformément à ω_d) et telle qu'on puisse passer d'un couple à l'autre par une chaîne intermédiaire de couples dont deux consécutifs ont une composante définie par une adresse commune. C'est par exemple le cas de la suite des adresses [(ligne, colonne)]: $\{(2,5),(1,2),(1,3),(3,4)\}$.

En effet, il est clair que l'exploration de toutes les solutions possibles, où à chaque pas on procède à la résolution d'un couple extrémal du tableau résultant du pas précédent, se fait au moyen d'une structure de pile.

Considérons l'exemple du tableau ci-dessus considéré de type
(100,70,10,10,7,1;80,50,50,8,3,3,2,2). **(10)**

La suite des couples extrémaux -ordonnés selon ω_d- est :
(7,8),(10,8),(70,80),(100,80).

La suite des adresses associées (numéro de ligne, numéro de colonne) est :
(5,4),(3,4) ou (4,4), (2,1), (1,1).

Nous allons également noter le type (10) sous la forme suivante :

$$t^{(0)}=\begin{Bmatrix} 100 & 70 & 10 & 10 & 7 & 1 & & \\ 80 & 50 & 50 & 8 & 3 & 3 & 2 & 2 \end{Bmatrix}, \quad \textbf{(11)}$$

où la première ligne (resp. seconde) correspond à la marge ligne (resp. colonne).

La résolution d'un couple (a_i,b_j) sera notée $r(a_i,b_j)$.$r(a_i,b_j)$ sera également considérée comme une transformation du type du tableau, résultant du déchargement de a_i dans j (resp. de b_j dans i) si $a_i<b_j$ (resp. $a_i>b_j$). On pourra noter :

$$t^{(l)}=r(a_i,b_j)\left[t^{(l-1)}\right], \quad \textbf{(12)}$$

où $t^{(l)}$ est le type du tableau obtenu après la l-ème résolution qui est $r(a_i,b_j)$.

Cette résolution incrémentera la somme des carrés des parts installées dans le tableau de contingence. On posera

$$r(a_i,b_j)\left[S^{(\ell-1)}\right]=S^{(\ell-1)}+\left[\min(a_i,b_j)\right]^2=S^{(\ell)}. \quad \textbf{(13)}$$

Le numéro d'ordre d'une résolution est compté à partir du type initial du tableau [cf.(**11**)] ; de sorte qu'on compte au plus (I+J-1) résolutions.

Dans l'exemple que nous allons traiter, on commence par empiler (10,8) et (7,8)

$$r(7,8)\left[t^{(0)}\right]=\left\{\begin{matrix}100 & 70 & 10 & 10 & 1 & & & \\ 80 & 50 & 50 & 3 & 3 & 2 & 2 & 1\end{matrix}\right\}=t^{(1)}$$

$$r(7,8)\left[S^{(0)}\right]=7^2=S^{(1)}.$$

L'algorithme profite de ce qui va être démontré au paragraphe III.2. et que nous avons déjà mentionné ci-dessus : si une même part marginale se retrouve des deux côtés, la solution optimale passe nécessairement par la résolution du couple concerné.

Ainsi, on poursuit nécessairement par la résolution du couple (1,1) :

$$r(1,1)\left[t^{(1)}\right]=\left\{\begin{matrix}100 & 70 & 10 & 10 & & & \\ 80 & 50 & 50 & 3 & 3 & 2 & 2\end{matrix}\right\}=t^{(2)}.$$

$$r(1,1)\left[S^{(1)}\right]=7^2+1^2=S^{(2)}.$$

En notant par "ce" l'ensemble des couples extrémaux, on écrit :

$$ce\left[t^{(2)}\right]=\{(10,3),(70,80),(100,80)\}$$

On n'empile que (10,3) qui se trouve déconnecté du reste

$$r(10,3)\left[t^{(2)}\right]=\left\{\begin{matrix}100 & 70 & 10 & 7 & & \\ 80 & 50 & 50 & 3 & 2 & 2\end{matrix}\right\}=t^{(3)}$$

$$r(10,3)\left[S^{(2)}\right]=7^2+1^2+3^2=S^{(3)}.$$

$$ce\left[t^{(3)}\right]=\{(7,3),(10,3),(70,80),(100,80)\}$$

On empile (10,3) et (7,3) qui correspondent respectivement aux adresses (3,4) et (4,4).

$$r(7,3)\left[t^{(3)}\right]=\begin{Bmatrix}100 & 70 & 10 & 4 & \\ 80 & 50 & 50 & 2 & 2\end{Bmatrix}=t^{(4)}$$

$$r(7,3)\left[S^{(3)}\right]=7^2+1^2+3^2+3^2=S^{(4)}$$

$$ce\left[t^{(4)}\right]=\{(4,2),(10,2),(70,80),(100,80)\}$$

On empile (10,2) et (4,2)

$$r(4,2)\left[t^{(4)}\right]=\begin{Bmatrix}100 & 70 & 10 & 2\\ 80 & 50 & 50 & 2\end{Bmatrix}=t^{(5)}$$

$$r(4,2)\left[S^{(4)}\right]=7^2+1^2+3^2+3^2+2^2=S^{(5)}$$

Nécessairement

$$r(2,2)\left[t^{(5)}\right]=\begin{Bmatrix}100 & 70 & 10\\ 80 & 50 & 50\end{Bmatrix}=t^{(6)}$$

$$r(2,2)\left[S^{(5)}\right]=7^2+1^2+3^2+3^2+2^2+2^2=S^{(6)}.$$

On empile (100,80) et (70,80) qui correspondent à $ce\left[t^{(6)}\right]$.

$$r(70,80)\left[t^{(6)}\right]=\begin{Bmatrix}100 & 10 & \\ 50 & 50 & 10\end{Bmatrix}=t^{(7)}$$

$$r(70,80)\left[S^{(6)}\right]=7^2+1^2+3^2+2^2+2^2+70^2=S^{(7)}$$
$$+3^2$$

Nécessairement

$$r(10,10)\left[t^{(7)}\right]=\begin{Bmatrix}100 & \\ 50 & 50\end{Bmatrix}=t^{(8)}$$

$$r(10,10)\left[S^{(7)}\right]=7^2+1^2+3^2+3^2+2^2+2^2+70^2+10^2$$
$$=S^{(8)}$$

Nécessairement

$$r(100,50)\left[t^{(8)}\right]=\begin{Bmatrix}50\\ 50\end{Bmatrix}=t^{(9)}$$

$$r(100,50)\left[S^{(8)}\right]=7^2+1^2+3^2+3^2+2^2+2^2+70^2+10^2+50^2=S^{(9)}$$

$$r(50,50)\left[t^{(9)}\right]=\{\ \}=t^{(10)}$$

$$r(50,50)\left[S^{(9)}\right]=7^2+1^2+3^2+3^2+2^2+2^2+70^2+10^2+50^2+50^2=S^{(10)}$$

On obtient ainsi la configuration T_5 (ci-dessus) pour laquelle $s(T_5)=S^{(10)}=10076$.

En remontant au dernier couple empilé (100,80) et en redescendant, on obtient une configuration T_8 pour laquelle

$$s(T_8)= 7^2+1^2+3^2+3^2+2^2+2^2+80^2+50^2+20^2+20^2+10^2= 9876.$$

Il faut remonter au couple empilé (10,8) et redescendre à partir de $t^{(6)}$ par le même chemin pour découvrir la solution optimale qui correspond à T_7 pour lequel $s(T_7)=1088$.

Signalons que pour ce type de tableau [cf. (10)], la borne "analytique" vaut 13.127,... et que $\min(\sum_i a_i^2, \sum_j b_j^2)= 11.490$.

III.2. Cas où une même part marginale se retrouve en ligne et en colonne

Un indice m évoquera ci-dessous le caractère maximal de s(T) ou de la configuration du tableau T.

THEOREME 1. Si dans un tableau T le contenu d'une marge ligne i se retrouve exactement dans le contenu d'une marge colonne j ($a_i=b_j=a$), la case (i,j) de la configuration optimale contient nécessairement la totalité a de la marge ligne (resp. colonne).

En d'autres termes, en dehors du contenu a de la case (i,j), la ligne i (resp. colonne j) ne contient plus que des zéros.

On ne restreint pas la généralité si on suppose (i,j)=(1,1). D'autre part, simplement pour fixer les idées et illustrer notre propos, nous allons prendre I=4 et J=5. Dans ces conditions, il s'agit de démontrer que

$$a^2+s_m\begin{bmatrix} a_2, & a_3, & a_4 & \\ b_2, & b_3, & b_4, & b_5 \end{bmatrix}$$

$$\geq (a-\alpha)^2+s_m\begin{bmatrix} \alpha & a_2 & a_3 & a_4 & \\ \alpha & b_2 & b_3 & b_4 & b_5 \end{bmatrix} \qquad (1)$$

Considérons la configuration optimale qui correspond au tableau du second membre :

$$\mathrm{Conf}_m\begin{bmatrix} \alpha & a_2 & a_3 & a_4 & \\ \alpha & b_2 & b_3 & b_4 & b_5 \end{bmatrix} \qquad (2)$$

dont nous désignons par $(\alpha_{1j}/j\geq 1)$ la première ligne et par $(\alpha_{i1}/i\geq 1)$, la première colonne.

Si $\alpha_{11} = \alpha$, on a

$$s_m\begin{bmatrix} \alpha & a_2 & a_3 & a_4 & \\ \alpha & b_2 & b_3 & b_4 & b_5 \end{bmatrix} = \alpha^2 + s_m\begin{bmatrix} a_2 & a_3 & a_4 & \\ b_2 & b_3 & b_4 & b_5 \end{bmatrix}$$

et la relation (1) est évidente de façon stricte si $\alpha \neq 0$.

$\alpha_{11} < \alpha$, alors il existe au moins un $j \geq 2$ (resp. $i \geq 2$), pour lequel α_{1j} (resp. α_{i1}) est différent de zéro. En effet,

$$\sum_i \alpha_{i1} = \sum_j \alpha_{1j} = \alpha.$$

On peut sans restreindre la généralité supposer que ce sont α_{21} et α_{12} qui sont en particulier différents de zéro et que

$$\alpha_{21} \geq \alpha_{12}$$

Nous supposons par ailleurs -dans le cadre d'un raisonnement par récurrence- que la propriété à démontrer est vraie pour tout entier représentant $\sum\{c_{ij}/1\leq i\leq I,\ 1\leq j\leq J\}$, inférieur ou égal à $(n-1)$.

	α	b_2	b_3	b_4	b_5
α	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}
a_2	α_{21}				
a_3	α_{31}				
a_4	α_{41}				

Dans ces conditions, nous remplaçons le tableau $T_m(\alpha)$ qui correspond au second membre de (1) par un tableau $T_0(\alpha-\alpha_{12})$ obtenu en vidant α_{12} dans la case (2,2) et en remplaçant α_{21} par $(\alpha_{21}-\alpha_{12})$. Cette opération préserve les contenus respectifs b_2 et a_2 de la deuxième composante de la marge ligne (resp. colonne). La perte maximale pouvant être encourue (dans le cas où $c_{22}=0$) est égale à

$$\alpha_{21}^2-(\alpha_{21}-\alpha_{12})^2$$

$$= \alpha_{12}(2\alpha_{21}-\alpha_{12}). \quad (3)$$

On a donc

$$s\left[T_o(\alpha-\alpha_{12})\right] \geq s\left[T_m(\alpha)\right]-\alpha_{12}(2\alpha_{21}-\alpha_{12})$$

A fortiori

$$s\left[T_m(\alpha-\alpha_{12})\right] \geq s\left[T_m(\alpha)\right]-\alpha_{12}(2\alpha_{21}-\alpha_{12}), \quad (4)$$

où $T_m(\alpha-\alpha_{12})$ est la structure optimale du tableau dont le type est défini par

$$\begin{bmatrix} (\alpha-\alpha_{12}) & a_2 & a_3 & a_4 & \\ (\alpha-\alpha_{12}) & b_2 & b_3 & b_4 & b_5 \end{bmatrix}.$$

Mais, compte tenu de l'hypothèse de récurrence

$$s\left[T_m(\alpha-\alpha_{12})\right]=(\alpha-\alpha_{12})^2+s_m\begin{bmatrix} a_2 & a_3 & a_4 & \\ b_2 & b_3 & b_4 & b_5 \end{bmatrix} \quad (5)$$

Ainsi, on obtient à partir de (4)

$$a^2+s_m\begin{bmatrix} a_2 & a_3 & a_4 & \\ b_2 & b_3 & b_4 & b_5 \end{bmatrix} \geq s(T_m(\alpha))+a^2-(\alpha-\alpha_{12})^2-\alpha_{12}(2\alpha_{21}-\alpha_{12}). \quad (6)$$

Pour établir le résultat (1), il suffit alors de vérifier que

$$a^2-(\alpha-\alpha_{12})^2-\alpha_{12}(2\alpha_{21}-\alpha_{12})>(a-\alpha)^2. \quad (7)$$

Simplifications effectuées, il revient de vérifier que

$$\alpha(a-\alpha)+\alpha_{12}(\alpha-\alpha_{21})>0 \quad (8)$$

ce qui est en effet le cas, puisque $\alpha \geq \alpha_{21}$ et que $a>\alpha$.

Quant à l'hypothèse de récurrence, il n'y a rien à démontrer pour n=1 et elle peut être directement vérifiée pour les premières valeurs de n. Le résultat du théorème est donc complètement acquis.

III.3. Cas d'un seul couple extrémal

Nous allons dans l'ordre traiter les cas suivants :

1 : I=2 et J=2
2 : I=2 et J quelconque.
3 : I quelconque et J=2
4 : I et J quelconques.

<u>**PROPRIETE 2.**</u> Dans une table 2x2 de $\mathcal{C}$, on a nécessairement :

$a_1>b_1\geqslant b_2>a_2$. (1)

De plus, le couple extrémal est défini par (a_1,b_1).

En effet, puisque la table appartient à $\mathcal{C}$, on a déjà :
$a_1\geqslant a_2$, $b_1\geqslant b_2$ et $a_1>b_1$. (2)

Nous allons raisonner relativement à la position de a_2, compte tenu de la suite déjà acquise des inégalités :

$a_1>b_1\geqslant b_2$. (3)

On ne peut avoir
$a_1\geqslant a_2>b_1\geqslant b_2$, (4)
car alors, on aurait
$(a_1+a_2)>(b_1+b_2)$. (5)

On ne peut non plus -car on serait conduit à la même conclusion (5)- avoir

$a_1>b_1\geqslant a_2\geqslant b_2$. (6)

On a donc
$a_1>b_1\geqslant b_2>a_2$.

Dans ces conditions, (a_1+b_1) est trivialement la plus grande somme possible entre les contenus respectifs d'une marge ligne et d'une marge colonne.

D'autre part, on a tout aussi trivialement
$(a_1-b_1)=(b_2-a_2)<(a_1-b_2)=(b_1-a_2)$. (7)

Le cas où l'inégalité de (7) se réduit à une égalité n'est pas à prendre en considération, puisqu'il conduit à

$$a_1=a_2=b_1=b_2.$$

THEOREME 3. Pour une table T, 2x2 de $\mathcal{C}$, la configuration maximale pour s(T) est définie par le tableau suivant :

	b_1	b_2
a_1	b_1	a_1-b_1
a_2	0	a_2

c'est-à-dire, par le déchargement de b_1 dans la première ligne.

Pour la précédente configuration que nous noterons (0), on a
$$s_o(T)=b_1^2+(a_1-b_1)^2+a_2^2$$
$$=a_1^2+a_2^2+2b_1(a_1-b_1). \quad (1)$$

Considérons à présent une configuration générale, notée (x), où b_1 se répartit en (b_1-x) dans la première ligne et en x dans la deuxième ligne :

	b_1	b_2
a_1	b_1-x	a_1-b_1+x
a_2	x	a_2-x

On a, compte tenu de (1) de la propriété ci-dessus,
$0<x<a_2$. (2)

Dans ce dernier tableau, on a
$$s_x(T)=x^2+(b_1-x)^2+(a_1-b_1+x)^2+(a_2-x)^2$$
$$=4x^2+2x(a_1-a_2-2b_1)+b_1^2+(a_1-b_1)^2+a_2^2. \quad (3)$$

Ainsi,

$$s_x(T)-s_o(T)=2x\left[2x+(a_1-a_2-2b_1)\right], \quad (4)$$

dont le signe est celui de

$$\left[2x+(a_1-a_2-2b_1)\right],$$

en raison de (2), cette dernière quantité reste inférieure à

$$(a_1+a_2-2b_1)=(n-2b_1) \quad (5)$$

et

$$(n-2b_1)\leqslant 0,$$

en raison de (1) de la propriété ci-dessus, CQFD.

THEOREME 4. Pour une table T, 2xJ de $\mathcal{T}$, pour laquelle le seul couple extrémal est (a_1,b_1), la solution maximale pour s(T) passe nécessairement par le déchargement total de b_1 dans la première ligne.

Considérons en effet la configuration suivante où, pour fixer les idées mais sans nullement restreindre la généralité, nous avons pris J=4.

	b_1	b_2	b_3	b_4
a_1	b_1-x	c_{12}	c_{13}	c_{14}
a_2	x	c_{22}	c_{23}	c_{24}

Remarquons qu'on a nécessairement

$$a_1>b_1>a_2, \quad (1)$$

en effet, si on avait

$$a_1\geqslant a_2>b_1,$$

(a_1,b_1) n'aurait pas été le seul élément extrémal. Ainsi,

$$x\leqslant a_2<b_1$$

Deux cas de figure peuvent se présenter :

$x\leqslant(b_1-x)$ ou bien $x>(b_1-x)$. (2)

Nous allons montrer que dans le premier cas -où $x\leqslant(b_1-x)$- si x n'est pas nul, on peut trouver une opération élémentaire, transformant x en (x-1) et (b_1-x) en (b_1+1-x), qui augmente s(T).

Remarquons qu'on ne peut avoir
$c_{2j}=0$ pour tout $j=2,\dots,J$.

En effet, a_1 ne peut se réduire à (b_1-x) qui est inférieur ou égal à b_1, alors que par hypothèse, $a_1>b_1$.

Maintenant, on peut même supposer sans restreindre la généralité que $c_{2j}\neq 0$ pour tout $j=2,3,\dots,J$. En effet, si $c_{1j_0}=0$ pour un $j_0\geqslant 2$, nous supprimons la colonne j_0, remplaçons a_2 par $(a_2-b_{j_0})$ et le problème se repose exactement dans les mêmes termes avec une colonne en moins.

La relation $x\leqslant(b-x)$, implique que
$$\sum_{2\leqslant j\leqslant J} c_{1j}\leqslant \sum_{2\leqslant j\leqslant J} c_{2j}, \quad (3)$$

il en résulte qu'il existe au moins j_1 tel que

$$c_{1j_1}\leqslant c_{2j_1}$$

On a donc, compte tenu de ce qui a été supposé ci-dessus,

$$0<c_{1j_1}\leqslant c_{2j_1}$$

On ne restreint en rien la généralité si on suppose $j_1=2$. Ainsi,

$$0<c_{12}\leqslant c_{22}.$$

L'opération élémentaire va transformer le tableau figuré ci-dessus en le suivant :

	b_1	b_2	b_3	b_4
a_1	b_1-x+1	$c_{12}-1$	c_{13}	c_{14}
a_2	$(x-1)$	$c_{22}+1$	c_{23}	c_{24}

Or, si γ_1 et γ_2 sont deux entiers tels que $\gamma_1\geqslant\gamma_2$, la variation de la somme des carrés lorsqu'on remplace γ_1 (resp. γ_2) par (γ_1+1) [resp. par (γ_2-1)] est égale à

$$\left[(\gamma_1+1)^2+(\gamma_2-1)^2\right]-(\gamma_1^2+\gamma_2^2)$$

$$=2(\gamma_1-\gamma_2)+2.$$

Donc cette variation est positive et supérieure à 2. Ainsi, l'opération élémentaire ci-dessus considérée accroît au moins de 2, la contribution de chacune des deux premières colonnes.

Ainsi, la solution optimale ne peut correspondre à la répartition (b_1-x,x) -avec $x \leqslant (b_1-x)$- sur la suite des deux lignes, que si $x=0$.

Considérons maintenant le cas de figure où $x>(b_1-x)$. On a alors nécessairement

$$\sum_{2\leqslant j\leqslant J} c_{1j} > \sum_{2\leqslant j\leqslant J} c_{2j} \quad (4)$$

Dans ces conditions, il existe au moins un indice j_1 pour lequel

$$c_{1j_1} > c_{2j_1} \quad (5)$$

Si c_{2j_1} est strictement positif, l'opération élémentaire conduisant à la configuration suivante :

	b_1		b_{j1}	
a_1	b_1-x-1	....	$c_{1j_1}+1$	...
a_2	$x+1$	....	$c_{2j_1}-1$	...

augmente la contribution à $s(T)$ de chacune des colonnes 1 et j_1.

Cette opération élémentaire n'est pas possible si $x=a_2$; mais dans ce cas, la configuration devient :

	b_1	b_2	b_3	b_4	
a_1	b_1-a_2	b_2	b_3	b_4	(I)
a_2	a_2	0	0	0	

Comme dans l'étude du cas précédent, les colonnes j_1 pour lesquelles la relation (5) a lieu avec $c_{2j_1}=0$, peuvent être par suppression et remplacement de a_1 par $(a_1-b_{j_1})$, évacués du problème.

En conclusion, la solution optimale ne peut correspondre à la répartition (b_1-x,x) -avec $x>(b_1-x)$- sur la suite des deux lignes, que si $x=a_2$.

Il reste à montrer qu'entre la configuration (I) ci-dessus, pour laquelle $x=a_2$ et celle (II) la suivante, pour laquelle x=0, c'est celle (II) qui l'emporte ; mais pour un choix optimal du tableau résultant :

	b_1	b_2	b_3	b_4
a_1	b_1	c_{12}	c_{13}	c_{14}
a_2	0	c_{22}	c_{23}	c_{24}

(II)

	b_2	b_3	b_4
(a_1-b_1)	c_{12}	c_{13}	c_{14}
a_2	c_{22}	c_{23}	c_{24}

(III)

Plus précisément, en posant

$\alpha_{11}=(a_1-b_1)$ et $\beta_{12}=(b_1-a_2)$, **(6)**

il y a lieu de montrer que

$b_1^2+s_m\left[a_2,\alpha_{11}\ ;\ b_2,b_3,b_4\right] >$

$a_2^2+s_m\left[a_1\ ;\ b_2,b_3,b_4,\beta_{12}\right]$. **(7)**

Or, le second membre vaut

$a_2^2+b_2^2+b_3^2+b_4^2+\beta_{12}^2$.

Il s'agit donc de montrer que

$a_2^2+b_2^2+b_3^2+b_4^2+\beta_{12}^2<b_1^2+s_m\left[a_2,\alpha_{11}\ ;\ b_2,b_3,b_4\right]$. **(8)**

α_{11} représentant la plus petite différence, on suppose avoir $\alpha_{11}\leq b_J$ (J=4 dans notre exemple illustratif). Il suffira en fait de considérer dans notre démonstration

$\alpha_{11}\leq b_2$. **(9)**

Bien que cela n'intervienne pas directement, on peut noter qu'on a nécessairement

$$a_2 > b_2 \ ;$$

sinon, la relation

$$\alpha_{11} = (a_1 - b_1) < (b_2 - a_2),$$

aurait conduit à

$$a_1 + a_2 = n < b_1 + b_2 \ ;$$

de sorte qu'on peut compléter (1) et écrire

$$a_1 > b_1 > a_2 > b_2 \geq b_3 \geq b_4. \quad (10)$$

Relativement au tableau $\left[a_2, \alpha_{11} \ ; \ b_2, b_3, b_4\right]$, considérons, compte tenu de (9), la répartition suivante :

	b_2	b_3	b_4
α_{11}	α_{11}	0	0
a_2	$b_2 - \alpha_{11}$	b_3	b_4

La valeur de s associée à cette décomposition est :

$$(a_1 - b_1)^2 + \left[b_2 - (a_1 - b_1)\right]^2 + b_3^2 + b_4^2. \quad (11)$$

Cette quantité (11) est -par définition- inférieure ou égale à $s_m\left[a_2, \alpha_{11} \ ; \ b_2, b_3, b_4\right]$. En lui ajoutant b_1^2, on obtient :

$$b_1^2 + b_2^2 + b_3^2 + b_4^2 + 2(a_1 - b_1)^2 - 2(a_1 - b_1)b_2. \quad (12)$$

Il suffit de prouver que la différence entre (12) et le premier membre de (8), est positive. Cette différence se met sous la forme :

$$b_1^2 + 2(a_1 - b_1)(a_1 - b_1 - b_2) - a_2^2 - (b_1 - a_2)^2$$

$$= 2\left[a_2(b_1 - a_2) + (a_1 - b_1)(a_1 - b_1 - b_2)\right]. \quad (13)$$

Comme $(a_1 - b_1)$ représente la plus petite différence positive et que $b_1 > a_2$, l'expression (13) est supérieure à

$$2(a_1 - b_1)\left[(a_1 + a_2) - (b_1 + b_2)\right] > 0. \qquad \text{C.Q.F.D.}$$

THEOREME 5. Pour une table T, Ix2 de $\mathcal{T}$, la solution maximale pour s(T) passe nécessairement par le déchargement total de b_1 dans la première ligne, si $b_1 > b_2$.

Nous allons, simplement pour fixer les idées et illustrer la démonstration, supposer I=4. Nous allons dans ces conditions comparer les deux configurations suivantes :

	b_1	b_2
a_1	b_1	a_1-b_1
a_2	0	a_2
a_3	0	a_3
a_4	0	a_4

(I)

	b_1	b_2
a_1	b_1-x_1	$a_1-b_1+x_1$
a_2	x_{21}	a_2-x_{21}
a_3	x_{31}	$a_3-x_{\ 1}$
a_4	x_{41}	a_4-x_{41}

(II)

où x_1 est un entier positif inférieur à b_1 et où

$$x_1=x_{21}+x_{31}+x_{41}, \quad (1)$$

x_{21}, x_{31} et x_{41} étant trois entiers positifs ou nuls.

Etudions la variation du critère s lorsqu'on remplace la configuration (II) par la configuration (I). Pour cela, désignons par Δ_i (i=1,2,3 et 4) l'accroissement (positif ou négatif) dû à la i-ème ligne. On a

$$\Delta_1=2b_1x_1-x_1^2-2(a_1-b_1)x_1-x_1^2$$

$$=-2(x_1^2-2b_1x_1+a_1x_1)$$

$$=-2x_1(x_1-2b_1+a_1). \quad (2)$$

Au niveau de l'une des lignes i (i=2,3 ou 4), on a

$$\Delta_i=2x_{i1}\ (a_i-x_{i1}). \quad (3)$$

De sorte que

$$\Delta_2+\Delta_3+\Delta_4=2\ (a_2x_{21}+a_3x_{31}+a_4x_{41}-x_{21}^2-x_{31}^2-x_{41}^2). \quad (4)$$

Au facteur 2 près, la somme des Δ_i se met sous la forme suivante :

$$-x_1^2-x_{21}^2-x_{31}^2-x_{41}^2+(2b_1-a_1)x_1$$

$$+a_2x_{21}+a_3x_{31}+a_4x_{41}. \quad (5)$$

D'où

$$\Delta_1+\Delta_2+\Delta_3+\Delta_4=2\{(2b_1-a_1-x_1)x_1$$

$$+(a_2-x_{21})x_{21}+(a_3-x_{31})x_{31}+(a_4-x_{41})x_{41}\}. \quad (6)$$

On a nécessairement

$$\sum_{2\leqslant i\leqslant 4}(a_i-x_{i1})x_{i1}\geqslant 0$$

D'autre part, la valeur minimale de $(2b_1-a_1-x_1)$ correspond à la valeur maximale de x_1 qui est telle que $(a_1-b_1+x_1)$ maximum ; soit

$$a_1-b_1+x_1=b_2$$
$$x_1=(b_1+b_2-a_1)>0.$$

Pour une telle valeur de x_1,
$(2b_1-a_1-x_1)=(b_1-b_2)\geqslant 0$

Ainsi si $b_1>b_2$, l'accroissement du critère s est strictement positif.
C.Q.F.D.

__LEMME 6.__ Si c_{ij} est l'entier le plus grand de la ligne i et de la colonne j, relativement à la configuration optimale d'un tableau T, alors c_{ij} est nécessairement égal au contenu a_i (resp.b_j) de la marge colonne (resp. ligne) si $a_i\leqslant b_j$ (resp.$b_j<a_i$).

En d'autres termes, en dehors du contenu a_i (si $a_i\leqslant b_j$) [resp. b_j (si $b_j<a_i$)], la ligne i (resp. colonne j), ne contient plus que des zéros.

On ne restreint en rien la généralité si on suppose i=j=1. Nous allons alors montrer que si c_{11} est strictement plus petit que $\min(a_1,b_1)$ pour une configuration donnée du tableau T, il est toujours possible de trouver une autre configuration qui augmente c_{11} et qui accroît s(T).

En effet, il existerait nécessairement $c_{i_o 1} \neq 0$ et $c_{1j_o} \neq 0$ tels que :

$$\max(c_{i_o 1}, c_{1j_o}) \leq c_{11}$$

On ne restreint en rien la généralité si on suppose que

$$c_{i_o 1} \geq c_{1j_o} > 0$$

et on a bien sûr :

$$0 \leq c_{i_o 1} - c_{1j_o} < c_{11}$$

Autour des quatre cases $(1,1),(1,j_o),(i_o 1)$ et (i_o,j_o), considérons la nouvelle configuration suivante qui ne perturbe en rien les contenus des marges :

	1	j_o
1	$(c_{11}+c_{1j_o})$	0
i_o	$(c_{i_o 1}-c_{1j_o})$	$(c_{i_o j_o}+c_{1j_o})$

Puisque $c_{i_o 1} \leq c_{11}$, il est clair que la transformation apporte en ce qui concerne la contribution de la colonne 1, un incrément strictement positif et d'ailleurs égal à

$$2\, c_{1j_o}\left[c_{1j_o}+(c_{11}-c_{i_o 1})\right].$$

D'autre part, il est clair que la nouvelle colonne j_o apporte à s(T) un incrément positif ou nul.

Ce lemme montre en particulier que -pour la configuration optimale- la plus grande valeur rencontrée dans le tableau est celle de la marge ligne ou colonne qui correspond à la case contenant cette plus grande valeur.

LEMME 7. T est une table de la famille $\mathcal{T}$ pour laquelle (a_1,b_1) est le seul élément extrémal. Si on a

$$a_1 \geq b_2+\ldots+b_J$$

et

$$(a_1-b_1) \leq b_J,$$

alors, la solution en équerre (où $c_{ij}=0$ dès lors que $\min(i,j)\geqslant 2$) ne peut être optimale.

Nous allons établir ce résultat par récurrence sur le nombre J de colonnes.

Nous avons déjà et de façon générale montré (cf. théorème 5 ci-dessus) que pour une telle table T de dimension Ix2 de $\mathcal{T}_0$, la solution maximale passe nécessairement par le déchargement total de b_1 dans la première ligne. Ce qui entraîne que $c_{i2}=a_i$ pour $i\geqslant 2$ et qui vérifie l'énoncé du lemme.

La propriété est dans ces conditions supposée vraie pour J colonnes (ou moins) et nous allons l'établir pour (J+1) colonnes. Dans l'illustration considérée ci-dessous pour fixer les idées, nous prenons I=3 et J=3. Nous allons plus précisément supposer selon l'hypothèse de récurrence qu'il y a une solution meilleure que celle en équerre, laquelle suppose d'une part le déchargement du reste (a_1-b_1) dans la dernière colonne J.

Considérons la solution en équerre pour une table T de dimension Ix(J+1) et répondant aux conditions du lemme :

	b_1	b_2	b_3	b_4
a_1	e_{11}	b_2	b_3	b_4
a_2	a_2	0	0	0
a_3	a_3	0	0	0

(I)

Considérons pour une telle répartition la fusion des deux dernières colonnes J et (J+1), on obtient :

	b_1	b_2	(b_3+b_4)
a_1	e_{11}	b_2	(b_3+b_4)
a_2	a_2	0	0
a_3	a_3	0	0

(II)

Les marges de la nouvelle table continuent de répondre et a fortiori aux conditions de l'énoncé du lemme puisque

$$a_1 \geq b_2+\ldots+b_{(J-1)}+\left[b_J+b_{(J+1)}\right]$$

$$(a_1-b_1) \leq b_{(J+1)} \leq \min\left[b_{(J-1)},(b_J+b_{(J+1)})\right] \quad (1)$$

Nous allons envisager les deux cas selon que le dernier membre est égal à $(b_J+b_{(J+1)})$ ou à $b_{(J-1)}$; la dernière colonne correspondant à la plus petite valeur.

Dans le premier cas, compte tenu de l'hypothèse de récurrence, on a relativement à la nouvelle table, une solution meilleure que celle (II) en équerre, laquelle correspond au déversement de b_1 dans la première ligne et de (a_1-b_1) dans la dernière colonne.

	b_1	b_2	(b_3+b_4)
a_1	b_1	0	(a_1-b_1)
a_2	0	c_{22}	C_{23}
a_3	0	c_{32}	C_{33}

(III)

Dans la configuration telle que (II) éclatons la dernière colonne en deux colonnes de marges respectives b_J et $b_{(J+1)}$ pour retrouver la configuration (I). La perte qui en résulte pour s(T) est de $2\ b_J \times b_{(J+1)}$.

Maintenant, à partir d'une configuration telle que (III), créons une colonne de marge b_J et une colonne de marge $b_{(J+1)}$, dont la première ligne contient (a_1-b_1). C'est en effet possible puisque par hypothèse $(a_1-b_1) \leq b_{(J+1)}$. La perte maximale encourue est inférieure à $2b_J\left[b_{(J+1)}-(a_1-b_1)\right]$

En effet, en posant

$$C_{2J}=\alpha_2 b_J+\beta_2\left[b_{(J+1)}-(a_1-b_1)\right]$$

$$C_{3J}=\alpha_3 b_J+\beta_3\left[b_{(J+1)}-(a_1-b_1)\right]$$

....

$$C_{IJ}=\alpha_I b_J+\beta_I\left[b_{(J+1)}-(a_1-b_1)\right], \quad (2)$$

où les coefficients α_i (resp. β_i) sont positifs et où

$$\alpha_2+\alpha_3+\ldots+\alpha_I=\beta_2+\beta_3+\ldots+\beta_I=1, \quad (3)$$

la perte encourue est exactement égale à

$$2(\alpha_2\beta_2+\ldots+\alpha_I\beta_I)b_J\left[b_{(J+1)}-(a_1-b_1)\right]. \quad \textbf{(4)}$$

Or

$$\alpha_2\beta_2+\ldots+\alpha_I\beta_I\leqslant(\alpha_2+\ldots+\alpha_I)(\beta_2+\ldots+\beta_I)=1. \quad \textbf{(5)}$$

Considérons à présent le cas où

$$\min\left[b_{(J-1)},(b_J+b_{(J+1)})\right]=b_{(J-1)}$$

On a bien sûr

$$(a_1-b_1)\leqslant b_J\leqslant b_{(J-1)}$$

et une solution meilleure que celle (II) en équerre (où pour illustrer, nous avons pris J=3) existe sous la forme :

$$\begin{array}{c|c|c|c|} & b_1 & b_2 & (b_3+b_4) \\ \hline a_1 & b_1 & (a_1-b_1) & 0 \\ \hline a_2 & 0 & c_{22} & C_{23} \\ \hline a_3 & 0 & c_{32} & C_{33} \\ \hline \end{array} \qquad \text{(IV)}$$

• Ici encore, à partir de la dernière colonne concaténée J, on constitue de l'une des façons possibles (il y en a toujours une) deux colonnes de marges respectives b_J et $b_{(J+1)}$. En raison d'un argument ci-dessus considéré (cf.(5)), la perte maximale est $2b_J x b_{(J+1)}$.

Dans ces conditions, la solution obtenue [par éclatement de la dernière colonne J en J et (J+1)] à partir de la configuration (III) ou bien de celle (IV) (avec comme nouvelles marges b_J et $b_{(J+1)}$) est meilleure que celle en équerre définie par la configuration (I). C.Q.F.D.

Pour être tout à fait complet par rapport à la démonstration ci-dessus, signalons que pour éclater la dernière colonne de la configuration (III) ou de celle (IV), il suffit de former la dernière colonne (J+1) en écrivant $b_{(J+1)}$ sous la forme

$$b_{(J+1)} = \sum_{2\leqslant i\leqslant I} c_{i(J+1)},$$

où **(6)**

$$c_{i(J+1)}\leqslant C_{iJ}$$

pour tout i=2,...,I.

Une telle écriture est toujours possible puisque dans chacun des deux cas -compte tenu notamment des conditions du lemme-

$$b_{(J+1)} \leq C_{2J}+\ldots+C_{IJ}. \quad (7)$$

THEOREME 8. Si (a_1,b_1) est le seul couple extrémal du tableau T de la famille $\mathcal{T}$, si de plus a_1 est plus grand ou égal que $(b_2+\ldots+b_J)$, alors la solution optimale passe nécessairement par la résolution du couple (a_1,b_1).

On commencera par remarquer que pour un tel tableau, on a nécessairement aussi

$$b_1 \geq a_2+\ldots+a_I. \quad (8)$$

Pour s'en rendre compte, il suffit en effet de considérer la configuration en équerre que ce tableau admet.

Nous allons démontrer cette propriété par récurrence sur le nombre total de lignes et de colonnes K=(I+J). Nous avons directement démontré la propriété dans le cas où (I+J)=4 (cf. Théorème 2 ci-dessus) et nous supposons qu'elle se trouve vérifiée pour tout tableau dont le nombre total K de lignes et de colonnes est tel que : $4 \leq K \leq (I+J-1)$. Il s'agit alors de la démontrer pour tout tableau T dont le nombre de lignes et de colonnes est (I+J). Soit T un tel tableau où on ne restreint en rien la généralité en supposant qu'il comporte I lignes, J colonnes et appartient à la famille $\mathcal{T}$. Ainsi, le couple (a_1,b_1) est extrémal.

Soit T_o la configuration optimale de ce tableau et soit $c_{i_oj_o}$ le plus grand entier de T_o. D'après le lemme 6,

$$c_{i_oj_o}=\min(a_{i_o},b_{j_o}) \quad (8)$$

et correspond au déchargement de a_{i_o} dans j_o (si $a_{i_o} \leq b_{j_o}$) ou bien de b_{j_o} dans i_o (si $b_{j_o} < a_{i_o}$).

Si $i_o \geq 2$ et $j_o \geq 2$ avec $a_{i_o} \leq b_{j_o}$. En enlevant la ligne i_o dont la seule case non nulle se trouve en colonne j_o (et contient a_{i_o}), on se retrouve devant un tableau à (I-1) lignes et J colonnes de type :

$$[a_1,a_2,\ldots,a_{(i_o-1)},a_{(i_o+1)},\ldots,a_I \; ; \; b_1,b_2,\ldots,b_{(j_o-1)},(b_{j_o}-a_{i_o}),b_{j_o+1},\ldots b_J],$$

où, manifestement, les conditions du théorème sont satisfaites. Pour ce dernier tableau, compte tenu de l'hypothèse de récurrence, la solution optimale passe par la résolution du couple (a_1,b_1) ; c'est-à-dire que la case (1,1) contient b_1. D'autre part, il est clair que si le tableau T (à I lignes et J colonnes) présente une configuration optimale, le fait d'enlever une ligne (resp. colonne), en modifiant en conséquence les marges colonnes (resp. lignes), préserve une configuration optimale pour le sous tableau. Car si on pouvait faire mieux au niveau du sous tableau, en remettant la ligne (resp. colonne) enlevée, on obtiendrait une contradiction quant au caractère optimal de la configuration de départ.

Le cas où $i_o \geqslant 2$ et $j_o \geqslant 2$ avec $b_{j_o} < a_{i_o}$ se traite exactement de la même façon, les rôles de la ligne i_o et de la colonne j_o étant intervertis.

Ainsi, la plus grande valeur de T_o, $c_{i_oj_o}$ ne peut se trouver dans la région interne du tableau définie par $\{(i,j)/i \geqslant 2$ et $j \geqslant 2\}$; il y aurait en effet contradiction avec le fait que c'est b_1 la plus grande valeur pouvant être rencontrée dans T_o (en vertu de l'hypothèse de récurrence). $c_{i_oj_o}$ se trouve soit sur la première ligne, soit sur la première colonne. Si elle se trouve à l'intersection c'est qu'elle est égale à b_1 et il n'y a plus rien à démontrer.

Imaginons alors que cette plus grande valeur se trouve sur la première ligne et qu'il s'agit alors nécessairement d'un b_j $(a_1 \geqslant b_2 + \ldots + b_J)$. On peut poser :

$$c_{i_oj_o} = c_{1j_o} = b_{j_o}$$

La résolution suivante, ni l'une quelconque de celles qui suivent, ne peut concerner un couple vierge (a_{i_1}, b_{j_1}) dans la région interne du tableau $(i_1 \geqslant 2,\ j_1 \geqslant 2)$. Car en effet, en faisant marche arrière et en reprenant la situation de départ, on arrive à une contradiction en ôtant la ligne i_1 (resp.j_1) si $a_{i_1} \leqslant b_{j_1}$ (resp. $b_{j_1} < a_{i_1}$), puisque le tableau ainsi délesté doit inclure b_1 dans sa case (1,1), en vertu de l'hypothèse de récurrence.

Imaginons qu'à une certaine étape, un certain nombre de b_j (resp.a_i) se trouvent résolus dans la première ligne (resp. colonne) dans le segment $i>1$ (resp.$j>1$) à partir de la détermination à chaque fois du plus grand entier se trouvant dans le tableau, comme il est indiqué dans l'exemple suivant :

	$\frac{5}{17}$	$\frac{0}{5}$	$\frac{0}{5}$	4	4
8‖18		5	5		
0‖6	6	0	0	0	0
0‖6	6	0	0	0	0
5		0	0		

(I)

Cet exemple répond bien aux conditions du théorème. Nous avons marqué la partie remplie du tableau, ainsi que -au delà des marges- ce qui reste à distribuer.

Ou bien le processus se poursuit avec le chargement total des sections $i>1$, de la première ligne (resp.$j>1$, de la première colonne) avant le chargement -avec ce qui reste- de la case (1,1), ou bien c'est la case (1,1) qui se charge au mieux, comme dans le tableau suivant qui se déduit du précédent (I) ci-dessus.

	$\frac{0}{17}$	$\frac{0}{5}$	$\frac{0}{5}$	4	4
3‖18	5	5	5		
0‖6	6	0	0	0	0
0‖6	6	0	0	0	0
5	0	0	0		

(II)

Dans le premier cas on a la configuration en équerre et le lemme 7 ci-dessus, nous montre que dans les conditions du théorème, la solution en équerre ne peut être optimale si $(a_1-b_1)<b_J$. Maintenant, si $(a_1-b_1)>b_J$, en enlevant la dernière colonne, on se retrouve devant le type suivant du tableau $[(a_1-b_J),a_2,\dots,a_I;b_1,b_2,\dots,b_{(J-1)}]$ qui continue à remplir les conditions du théorème et pour lequel la plus petite différence entre une marge libre et une marge colonne diminue en passant de (a_1-b_1) à $|(a_1-b_1)-b_J|>0$.

D'autre part, on a nécessairement $a_1>b_1>a_2$, sinon, (a_1-b_1) n'aurait pas constitué la plus petite différence. On a nécessairement

$$(b_1-a_2)>(a_1-b_1).$$

Si $[(a_1-b_1)-b_J]>0$, c'est que -et c'est équivalent- $(a_1-b_J)>b_1$.

Donc (a_1-b_J) -qui constitue la nouvelle marge de la première ligne- et b_1 conduisent au couple qui réalise la plus grande somme. $[(a_1-b_J),b_1]$ est donc le seul élément extrémal du nouveau tableau.

Si $[(a_1-b_1)-b_J]$ est inférieur à $b_{(J-1)}$, l'application du lemme 7 montre que la configuration en équerre résultante à I lignes et (J-1) colonnes, ne peut être optimale. Si $[(a_1-b_1)-b_J]$ est strictement supérieur à $b_{(J-1)}$, on procède comme avant en enlevant la $(J-1)^e$colonne... Fatalement, on va arriver à une différence entre la première marge ligne et la première marge colonne inférieure à la dernière marge colonne, puisque $a_1<b_1+b_2+...+b_J=n$.

Le dernier (mais non le moindre) problème qui se pose maintenant est celui où au cours du processus, on découvre que le plus grand entier se trouve -comme il est indiqué dans le tableau (II) ci-dessus- dans la case (1,1). En désignant par $\sum'_j$ (resp.$\sum'_i$) une somme partielle portée sur $\{j/j\geq 2\}$ (resp.$\{i/i\geq 2\}$), ce contenu de la case (1,1) va alors se présenter soit sous la forme

$$[a_1-\sum'_j b_j], \quad (9)$$

soit sous la forme

$$[b_1-\sum'_i a_i]. \quad (10)$$

Dans le premier (resp. second) cas la première marge ligne (resp. colonne) se vide et les dernières marges colonnes (resp. lignes) [celles non impliquées dans la somme $\sum'_j$ (resp. $\sum'_i$)] doivent se répartir en dehors de la première ligne (resp. colonne). Ainsi, dans l'exemple ci-dessus où

$$c_{11}=b_1-(a_2+a_3),$$

la dernière ligne doit se répartir ailleurs que dans les colonnes 1 et celles impliquées dans $\sum'_j$.

Continuons la preuve en nous appuyant sur l'exemple qui n'a rien de particulier. Le prochain plus grand entier va se trouver dans l'une des colonnes non encore touchées. S'il se trouve en dehors de la première ligne, il correspond nécessairement comme dans l'exemple ci-dessous à la résolution d'un couple (a_i,b_j) dans la région $\{(i,j)/i\geqslant 2, j\geqslant 2\}$ du tableau. Mais, c'est impossible compte tenu de l'hypothèse de récurrence.

	17	5	5	4	4
18	5	5	5	0	3
6	6	0	0	0	0
6	6	0	0	0	0
5	0	0	0	4	1

(III)

Si le prochain plus grand entier se trouve sur la première ligne, comme dans l'exemple suivant. S'il s'agit d'un b_j, on se pose la même question. De sorte qu'on peut supposer que tous les b_j intégrés au niveau de la première ligne ont déjà été pris en considération. Il s'agit donc du reste de la forme $(a_1-\sum_j^! b_j)$ dans la dernière colonne disponible dont nous noterons l'indice j_1. Mais alors les marges lignes laissées lors de la prise en compte du chargement de la case (1,1) au moyen de $(b_1-\sum_i^! a_i)$ doivent nécessairement se résoudre totalement -comme dans l'exemple ci-dessous où il s'agit de la quatrième ligne- au niveau de la colonne j_1. On retrouve une fois de plus une impossibilité compte tenu de l'hypothèse de récurrence.

	17	5	5	4
18	5	5	5	3
6	6	0	0	0
6	6	0	0	0
1	0	0	0	1

(IV) C.Q.F.D.

THEOREME PRINCIPAL 9. Si (a_1,b_1) est le seul couple extrémal du tableau T de la famille $\mathcal{T}_0$, alors la configuration optimale du tableau passe nécessairement par la résolution du couple (a_1,b_1) (déchargement de b_1 dans la première ligne).

Ce résultat qui utilisera pour sa démonstration le théorème 8 sera également établi par récurrence et de la même manière. Le théorème a déjà été établi dans le cas où $(I+J)=4$ (cf. théorème 2 ci-dessus). Nous supposons qu'il est vrai pour tout tableau dont le nombre total K de lignes et de colonnes est tel que : $4 \leqslant K \leqslant (I+J-1)$. Il s'agit alors de le démontrer pour tout tableau T dont le nombre de lignes et de colonnes est $(I+J)$.

On peut comme ci-dessus supposer que

$a_1 > b_1 > a_2$. (1)

Car si on avait $(a_1 > a_2 > b_1$, $(a_1 - b_1)$ ne serait plus "la plus petite différence". D'autre part, si $a_1 = b_1$, le problème se trouve résolu au moyen du théorème 1 (cf.§III.2.). Enfin, nécessairement $(b_1 - a_2) \geqslant (a_1 - b_1)$, puisque (a_1, b_1) est le seul couple extrémal.

On peut toujours comme ci-dessus supposer que $(a_1 - b_1) < b_J$. Sinon, on aurait

$(a_1 - b_J) \geqslant b_1$

de sorte qu'en enlevant la dernière colonne de la configuration supposée optimale du tableau T : $Conf_m(T)$, la nouvelle marge ligne α_1 est telle que

$\alpha_1 \geqslant (a_1 - b_J)$.

Donc

$\alpha_1 \geqslant b_1 > a_2$

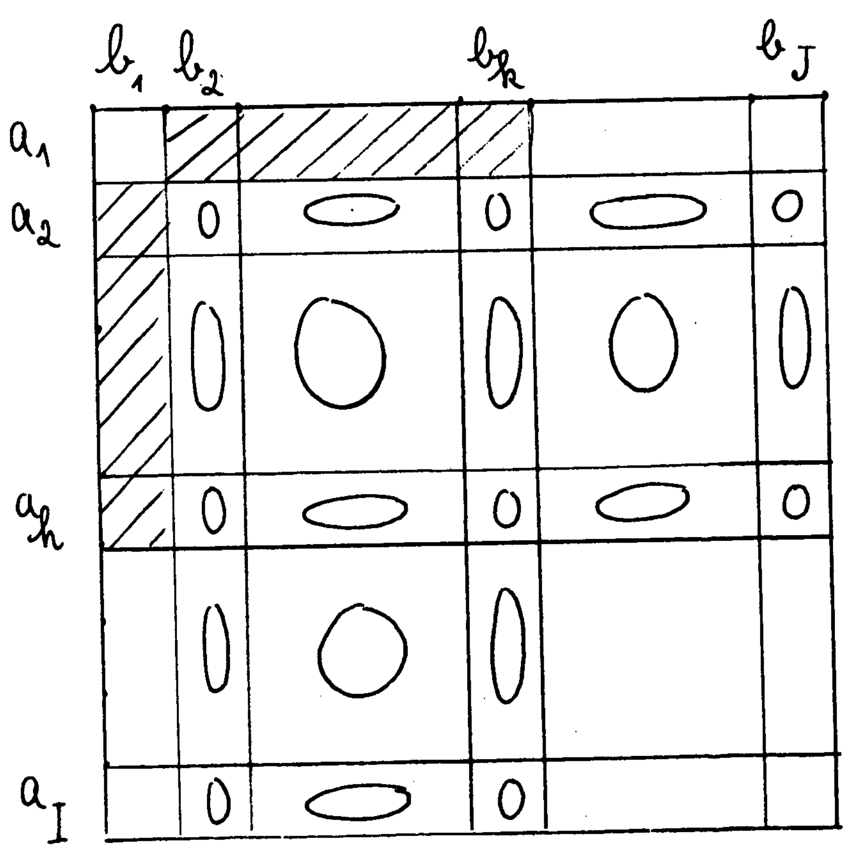
et (α_1, b_1) est le seul couple extrémal du tableau à I et (J-1) colonnes. Compte tenu de l'hypothèse de récurrence, pour la configuration optimale de ce dernier tableau, b_1 se trouve déchargé au niveau de la première ligne. Mais $Conf_m(T)$ comprend nécessairement la configuration optimale du sous tableau [comprenant les I lignes et les (J-1) premières colonnes] obtenu par suppression de la dernière colonne et réaménagement conséquent des marges lignes. Le théorème se trouve ainsi démontré si $(a_1 - b_1) \geqslant b_J$.

Continuons à appeler "intérieur" du tableau T la zone définie par $\{(i,j) / 1 < i \leqslant I$ et $1 < j \leqslant J\}$ où la première ligne et la première colonne sont omises. Considérons alors $Conf_m(T)$. Si cette configuration optimale comprend la résolution d'un seul couple (a_i, b_j) à l'intérieur, le théorème se trouve établi compte tenu de l'hypothèse de récurrence, puisque (a_1, b_1) reste le seul couple extrémal d'un tableau comprenant (I+J-1) lignes ou colonnes.

Nous allons chercher dans notre processus de démonstration à éviter cette situation -où il n'y a plus rien à démontrer- de la résolution d'un couple à l'intérieur.

On sait (lemme 6) que le plus grand entier du tableau correspond à la résolution d'un couple. Pour qu'il reste quelque chose à démontrer, cette résolution se fera au niveau du "bord" $\{ (i,j)/i=1,j>1 \text{ ou } j=1,i>1 \}$. Il s'agira donc soit d'un b_j déchargé au niveau de la première ligne ($j \geq 2$), soit d'un a_i déchargé au niveau de la première colonne ($i \geq 2$).

Imaginons qu'il y a eu en tout -en considérant à chaque fois le plus grand entier du tableau- (k-1) résolutions du type précédent au niveau de la première ligne et (h-1) au niveau de la première colonne. Pour fixer les idées -mais sans restreindre la généralité- on peut imaginer que les marges colonnes (resp. lignes) déchargées sont respectivement $b_2,b_3,\ldots b_k$ (resp. $a_2,a_3,\ldots,a_h$). On se retrouve alors devant la situation suivante où la partie hachurée indique les déchargements effectués. Nous allons imaginer la suite pas à pas de la construction de la configuration optimale où chaque pas correspond à la résolution nécessaire d'un couple du tableau délesté d'une ligne ou d'une colonne.



On peut déjà remarquer que si k=J (resp.h=I), alors nécessairement h=I (resp. k=J). Supposons dans ces conditions qu'on a $k<J$ et $h<I$ et posons

$$A=a_1-(b_2+...+b_k) \text{ et } B=b_1-(a_2+...+a_h). \quad (1)$$

Toujours et encore une fois pour qu'il reste à démontrer par rapport à l'hypothèse de récurrence, le prochain déchargement sera soit celui de A au niveau de la première ligne dans une colonne j_1, soit celui de B au niveau de la première colonne dans une ligne i_1. On ne restreint en rien la généralité si on suppose que c'est A qui se trouve déchargé dans la colonne (k+1). L'étape suivante doit nécessairement (sinon c'est fini) comprendre le déchargement de $[b_{(k+1)}-A]$ dans la ligne i_1 et on ne restreint pas la généralité si on suppose $i_1=(h+1)$. La suivante doit correspondre au déchargement de $a_{(h+1)}-(b_{(k+1)}-A)=A+a_{(h+1)}-b_{(k+1)}$ dans une colonne j_2 qu'on peut supposer (k+2). De même l'étape qui suit doit décharger le contenu de la marge colonne (k+2):$[b_{(k+2)}+b_{(k+1)}]-[A+a_{(h+1)}]$ dans une ligne i_2 qu'on peut supposer (h+2), etc...

En d'autres termes, pour éviter la résolution d'un couple à l'intérieur, chaque étape doit correspondre au déchargement de la marge qui a été touchée à l'étape qui précède. Mais, alors, en fin de parcours et fatalement, imaginons -sans restreindre la généralité- que la dernière marge touchée soit une marge ligne et qu'il s'agisse de la dernière marge ligne disponible alors qu'il reste encore au moins une colonne vierge. Les marges colonnes vierges restantes vont nécessairement être déchargées au niveau de cette dernière ligne disponible. Par conséquent, si $k<J$ et $h<I$, il y a nécessairement une résolution qui se produit à l'intérieur du tableau de la configuration optimale.

Le seul cas qui reste à envisager pour que la démonstration soit complète est celui où k=I et h=J. Mais ce cas a déjà été envisagé au niveau du théorème 8 précédent. C.Q.F.D.

III.4. Cas de plus d'un seul couple extrémal

Nous allons à présent considérer un ensemble de situations où il existe plus d'un seul couple extrémal et nous allons prouver que la configuration optimale du tableau de contingence comprend nécessairement la résolution d'un couple extrémal. On se situe bien sûr dans le cadre de la famille $\mathcal{T}$ de tableaux (cf. §III.1.), ce qui -encore une fois- ne restreint en rien la généralité. Ici, pour un tableau T de $\mathcal{T}$, (a_1,b_1) -qui réalise la plus grande somme- n'est pas l'unique couple extrémal ;

en d'autres termes, il existe (a_i,b_j) tel que $|a_i-b_j|<(a_1-b_1)$.

On désignera par $\mathcal{C}_1$ la sous famille de $\mathcal{C}$ pour laquelle il existe plus du seul couple extrémal (a_1,b_1).

THEOREME 10. Si pour le tableau T de $\mathcal{C}_1$, l'un des partages marginaux est plus fin que l'autre, la configuration optimale suppose nécessairement la résolution d'un couple extrémal au niveau de la première colonne.

Puisque -en vertu du théorème 1- on est dans le cadre $a_1>b_1$, c'est le partage $(b_1,b_2,\ldots,b_j,\ldots,b_J)$ qui est supposé plus fin que celui $(a_1, a_2,\ldots,a_i,\ldots,a_I)$. On sait que dans ce cas (cf.§II.3) la configuration optimale comprend les entiers b_j dans les cases du tableau.

Si cette configuration optimale comprend la résolution du couple (a_1,b_1), il n'y a plus rien à démontrer. Sinon, b_1 a nécessairement à être déchargé au niveau d'une ligne i qu'on notera i_o. Nous allons montrer que (a_{i_o},b_1) est nécessairement un couple extrémal.

Puisque $a_{i_o}<a_1$ [sinon la résolution de (a_{i_o},b_1) est équivalente à celle de (a_1,b_1)], on a

$$(a_{i_o}-b_1)<(a_1-b_1).$$

Maintenant, les seuls couples qui pourraient réaliser une somme plus grande que $(a_{i_o}+b_1)$ sont d'une part, les couples de la forme (a_i,b_1) où $i<i_o$. Mais pour ceux-là -compte tenu de la manière dont sont rangés les a_i- $(a_i-b_1)>(a_{i_o}-b_1)$ et (a_{i_o},b_1) reste un couple extrémal.

D'autre part, on peut avoir une somme plus grande pour un couple de la forme (a_{i_1},b_j) où $i_1<i_o$ et $j\geq 2$. Pour fixer les idées mais sans restreindre la généralité, on peut supposer que j=2.

Si $a_{i_1}>b_2$, on a
$(a_{i_1}-b_2)>(a_{i_o}-b_2)$ qui est positif.

En effet,

$a_{i_o} > b_1 \geq b_2$

et

$a_{i_o} < a_{i_1}$

Donc (a_{i_o}, b_1) reste extrémal.

Le dernier cas est celui où $b_2 > a_{i_1}$. Mais ce dernier est impossible. En effet, on aurait

$a_{i_o} > b_1 \geq b_2 > a_{i_1}$

et

$a_{i_o} + b_1 > a_{i_1} + b_2$, d'où la contradiction.

THEOREME 11. Dans le cas d'un tableau 2xJ de $\mathcal{O}_1$ [où (a_1,b_1) n'est pas le seul élément extrémal], la solution optimale passe nécessairement par la résolution de (a_1,b_1) ou de celle de (a_2,b_1).

Preuve

Nous allons établir ce résultat par récurrence sur le nombre de colonnes. Nous le supposons vrai jusqu'à (J-1) et il a naturellement été établi pour J=2 colonnes (cf. Théorème 3 § III.3.).

Nous allons considérer deux cas de figure. Le premier est celui où $a_1 \geq a_2 > b_1$ et le second est celui où $a_1 > b_1 > a_2$. Le cas où $a_2 = b_1$, où celui où $a_1 = b_1$ est résolu dans le cadre du théorème 1 ci-dessus.

Cas où $a_1 \geq a_2 > b_1$

Nous allons raisonner sur la plus grande valeur se trouvant dans le tableau. S'il s'agit de b_1, il n'y a plus rien à démontrer. S'il s'agit de b_{j_o} avec b_{j_o} strictement inférieur à b_1, on peut commencer par supposer pour fixer les idées et sans aucunement restreindre la généralité, que $j_o = 2$. Le cas où $b_{j_o} = b_1$ aurait résolu le problème moyennant la permutation des deux colonnes 1 et j_o. Il y a maintenant deux cas à considérer : le cas où la plus grande valeur du tableau b_2 $(b_1 > b_2)$ se trouve au niveau de la première ligne et le cas où b_2 se trouve au niveau de la deuxième ligne.

<u>Cas où b_2 est sur la 1ère ligne</u>

En enlevant la deuxième colonne, on se retrouve avec la configuration suivante où on fixe J=5.

$$[(a_1-b_2),a_2 \ ; \ b_1,b_3,b_4,b_5]$$

Si $(a_1-b_2) \geq b_1$, compte tenu de l'hypothèse de récurrence, on va nécessairement trouver b_1 au niveau de la première colonne. C'est impossible car la plus grande valeur du tableau initial est b_1 et non b_2.

Si maintenant, $a_2 > b_1 > (a_1-b_2)$.

Compte tenu de l'hypothèse de récurrence, on a soit la résolution de (a_2,b_1), soit celle de $[(a_1-b_2),b_1]$. La résolution du premier couple entraîne une contradiction. Celle du second couple conduit à la configuration suivante :

	b_1	b_2	b_3	b_4	b_5
a_1	a_1-b_2	b_2	0	0	0
a_2	$b_1+b_2-a_1$	0	b_3	b_4	b_5

dont une sous configuration compatible est la suivante :

	b_1	b_2
a_1	a_1-b_2	b_2
$(b_1+b_2)-a_1$	$b_1+b_2-a_1$	0

Or, pour ce tableau 2x2, une meilleure configuration se trouve définie par

	b_1	b_2
a_1	b_1	a_1-b_1
$(b_1+b_2)-a_1$	0	$b_1+b_2-a_1$

D'où la contradiction puisque b_2 ne pouvait être le plus grand entier du tableau optimal.

Cas où b_2 est sur la 2ème ligne

La démarche est la même que dans le précédent cas. On se retrouve devant la configuration

$$[a_1,(a_2-b_2);b_1,b_3,b_4,b_5]$$

où visiblement le seul cas à considérer est défini par la configuration suivante où $(a_2-b_2)<b_1$:

	b_1	b_2	b_3	b_4	b_5
a_1	$b_1+b_2-a_2$	0	b_3	b_4	b_5
a_2	a_2-b_2	b_2	0	0	0

dont une sous configuration compatible est définie par

	b_1	b_2
a_2	a_2-b_2	b_2
$(b_1+b_2-a_2)$	$b_1+b_2-a_2$	0

Mais alors, cette configuration peut avantageusement être remplacée par

	b_1	b_2
a_2	b_1	(a_2-b_1)
$(b_1+b_2-a_2)$	0	$b_1+b_2-a_2$

D'où la contradiction.

Cas où $a_1 > b_1 > a_2$.

Nous raisonnons toujours sur la plus grande valeur se trouvant dans le tableau. On reprend donc la même introduction que dans le cas où $a_1 \geq a_2 > b_1$ pour enclancher sur

Cas où b_2 est sur la 1ère ligne

On se retrouve face à la configuration suivante, après avoir enlevé la deuxième colonne :

$\left[(a_1-b_2),a_2;b_1,b_3,b_4,b_5\right]$.

L'hypothèse de récurrence étant vérifiée à ce niveau, on passe soit par la résolution de (a_2,b_1), soit par la résolution de $\left[(a_1-b_2),b_1\right]$. Dans le premier cas, il n'y a plus rien à démontrer. Dans le second cas, si $(a_1-b_2)>b_1$, on arrive à une contradiction puisqu'on suppose que b_2 est le plus grand entier trouvé dans le tableau. Il reste donc à examiner la situation suivante :

	b_1	b_2	b_3	b_4	b_5
a_1	(a_1-b_2)	b_2	0	0	0
a_2	$(b_1+b_2-a_1)$	0	b_3	b_4	b_5

dont une sous configuration compatible est définie par le tableau 2x2 :

	b_1	b_2
a_1	(a_1-b_2)	b_2
$(b_1+b_2-a_1)$	$b_1+b_2-a_1$	0

Cette configuration peut avantageusement être remplacée par

	b_1	b_2
a_1	b_1	(a_1-b_1)
$(b_1+b_2-a_1)$	0	$b_1+b_2-a_1$

,

d'où l'impossibilité.

Cas où b_2 est sur la 2ème ligne.

En enlevant la deuxième colonne, on se retrouve face à la configuration suivante :

$\left[a_1,(a_2-b_2);b_1,b_3,b_4,b_5\right]$

Compte tenu de l'hypothèse de récurrence, on passe soit par la résolution de (a_1,b_1), soit par la résolution de $\left[(a_2-b_2),b_1\right]$. On comprend qu'il nous reste à examiner cette dernière résolution qui conduit à la configuration suivante :

	b_1	b_2	b_3	b_4	b_5
a_1	$b_1+b_2-a_2$	0	b_3	b_4	b_5
a_2	a_2-b_2	b_2	0	0	0

,

où b_2 est supposé le plus grand. Par conséquent, on aurait :

$$b_2 \geqslant (a_2-b_2) \geqslant (b_1+b_2-a_2)$$

et (b_1-a_2) serait négatif. Ce qui est impossible compte tenu du cas où on se trouve. C.Q.F.D.

Remarques. Relativement au théorème précédent, il est important de noter que dans le cas où $a_1 \geqslant a_2 > b_1$, (a_2,b_1) est nécessairement un couple extrémal. En effet (a_1,b_j) ne peut être extrémal pour $j \geqslant 2$, puisque (a_1,b_j) se trouve alors dépassé par (a_1,b_1). D'autre part (a_2,b_j) ne peut être extrémal pour $j \geqslant 2$, puisque (a_2,b_j) se trouve alors dépassé par (a_2,b_1). Enfin, on a $(a_2-b_1) \leqslant (a_1-b_1)$. Maintenant, dans le cas où $a_1 > b_1 > a_2$, (a_2,b_1) est extrémal si $b_1 < n/2$. En effet, on est alors assuré que $(b_1-a_2)<(a_1-b_1)$ puisque $a_1+a_2=n$. Et, s'il y a un couple qui réalise une différence plus petite que (b_1-a_2), il est de la forme $(a_2,b_j)(j \geqslant 2)$ et pour ce dernier

$$(a_2+b_j)<(a_2+b_1)$$

Le théorème 5 ci-dessus concernant une table T de dimension Ix2, ne spécifie pas si (a_1,b_1) est le seul couple extrémal, on peut donc l'exprimer pour une table T Ix2 de $\mathcal{T}_1$.

La propriété suivante apparaîtra bien particulière, elle est en réalité susceptible d'une généralisation sur la position des résolutions et fait pressentir la nécessité de la résolution d'un couple extrémal dans la configuration optimale.

PROPRIETE 12. T est un tableau de la famille $\mathcal{T}_1$ avec au moins 3 colonnes. Le déchargement de b_2 au niveau de la première ligne ne peut conduire à une configuration optimale pour laquelle $c_{11} \geqslant (b_1-b_2)>0$. Si $c_{11}<(b_1-b_2)$, ce déchargement ne peut toujours pas conduire à la configuration optimale si (b_1-b_2) est supérieur à deux fois la différence entre la moyenne arithmétique et celle géométrique des deux nombres b_2 et $(a_2+b_2+b_3)$: $(b_1-b_2) \geqslant 2\{\frac{1}{2}\left[b_2+(a_2+b_2+b_3)\right]-\sqrt{b_2(a_2+b_2+b_3)}\}$

Preuve

Nous allons pour fixer les idées, mais sans aucunement restreindre la généralité, considérer I=3 et J=4. Nous aurons dans ces conditions à confronter les deux situations shématisées ci-dessous où nous comparons le déchargement de b_1 à celui de b_2 dans la première ligne. On considère dans chacun des cas la meilleure solution possible. Nous désignons par γ le contenu de la case (1,1) pour la seconde configuration.

	b_1	b_2	b_3	b_4
α_{11} a_1	b_1	.	.	.
a_2	0	.	.	.
a_3	0	.	.	.

	b_1	b_2	b_3	b_4
α_{12} a_1	γ	b_2	.	.
a_2	.	0	.	.
a_3	.	0	.	.

Nous notons $\alpha_{11}=(a_1-b_1)<\alpha_{12}=(a_1-b_2)$ et nous allons alors montrer que si $\gamma \geq (b_1-b_2)$, on a :

$$b_1^2+s_m\begin{bmatrix} a_2 & a_3 & \alpha_{11} \\ b_2 & b_3 & b_4 \end{bmatrix} \geq b_2^2+s_m\begin{bmatrix} a_2 & a_3 & \alpha_{12} \\ b_1 & b_3 & b_4 \end{bmatrix} \qquad (1)$$

A partir de la configuration optimale correspondant au second membre, nous allons former une configuration

$$\mathrm{Conf}\begin{bmatrix} a_2 & a_3 & \alpha_{11} \\ b_2 & b_3 & b_4 \end{bmatrix} \qquad (2)$$

et ce, en enlevant dans les cases de la première ligne -totalisant α_{12}- la quantité $(\alpha_{12}-\alpha_{11})$. Cet entier qui vaut

$$(\alpha_{12}-\alpha_{11})=(a_1-b_2)-(a_1-b_1)=(b_1-b_2)$$

est à enlever de la première colonne puisque le contenu γ de la case (1,1) est supérieur ou égal à (b_1-b_2).

La perte maximale pour s, encourue dans ce cas, résulterait du remplacement de b_1^2 par b_2^2 . Ainsi, pour la configuration (2) construite, on a

$$s \geqslant s_m\begin{bmatrix} a_2 & a_3 & \alpha_{12} \\ b_1 & b_3 & b_4 \end{bmatrix} - (b_1^2 - b_2^2). \quad (3)$$

Par conséquent,

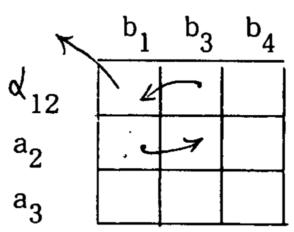
$$s_m\begin{bmatrix} a_2 & a_3 & \alpha_{11} \\ b_2 & b_3 & b_4 \end{bmatrix} \geqslant s_m\begin{bmatrix} a_2 & a_3 & \alpha_{12} \\ b_1 & b_3 & b_4 \end{bmatrix} - b_1^2 + b_2^2, \quad (4)$$

qui est exactement la relation (1) ci-dessus.

Nous allons maintenant considérer le cas où γ est strictement inférieur à (b_1-b_2) pour la configuration optimale obtenue après avoir placé b_2 sur la première ligne.

Pour former une configuration telle que (2), il va bien falloir déplacer dans la case (1,1), le complément $(\beta_{12}-\gamma)$ où nous avons noté $\beta_{12}=(b_1-b_2)$.

Si on considère la configuration optimale de $(a_2,a_2,\alpha_{12};b_1,b_3,b_4)$, argument du second membre de (1), la perte maximale encourue pour s résulte de la situation où b_3 (resp.a_2) se trouve sur la première ligne (resp. colonne) et où on effectue la transformation suivante portant sur $(\beta_{12}-\gamma)$.



La flèche sur le coin gauche haut indique -après la transformation- le dégagement de β_{12}.

Cette perte maximale pour $s_m(a_2,a_3,\alpha_{12};b_1,b_3,b_4)$ résulte des remplacements suivants :

$$\begin{aligned} b_3 &\longrightarrow b_3-(\beta_{12}-\gamma) \\ a_2 &\longrightarrow a_2-(\beta_{12}-\gamma) \\ \gamma &\longrightarrow 0 \\ 0 &\longrightarrow (\beta_{12}-\gamma) \end{aligned}$$

qui concernent respectivement les cases (1,2), (2,1) (1,1) et (2,2) (nouvelle numérotation).

La configuration résultante est du type (2) et on a

$$s_m\begin{bmatrix} a_2 & a_3 & \alpha_{12} \\ b_1 & b_3 & b_4 \end{bmatrix}+\{\left[b_3-(\beta_{12}-\gamma)\right]^2-b_3^2$$

$$+\left[a_2-(\beta_{12}-\gamma)\right]^2-a_2^2-\gamma^2+(\beta_{12}-\gamma)^2\}$$

$$\leqslant s_m\begin{bmatrix} a_2 & a_3 & \alpha_{11} \\ b_2 & b_3 & b_4 \end{bmatrix}. \quad (3)$$

Pour avoir la relation (1) ci-dessus, il suffit d'avoir

$$b_1^2-b_2^2+\left[b_3-(\beta_{12}-\gamma)\right]^2-b_3^2+\left[a_2-(\beta_{12}-\gamma)\right]^2$$

$$-a_2^2-\gamma^2+(\beta_{12}-\gamma)^2\geqslant 0. \quad (4)$$

Soit

$$\beta_{12}(b_1+b_2)-(\beta_{12}-\gamma)\left[2b_3-(\beta_{12}-\gamma)\right]$$

$$-(\beta_{12}-\gamma)\left[2a_2-(\beta_{12}-\gamma)\right]+\beta_{12}(\beta_{12}-2\gamma)\geqslant 0$$

qui se réduit à

$$\beta_{12}(b_1-\gamma)+(\beta_{12}-\gamma)\left[(\beta_{12}-\gamma)-a_2-b_3\right]\geqslant 0$$

et qu'on peut mettre sous la forme

$$\xi^2+\left[(b_1-b_2)-(a_2+b_3)\right]\xi+(b_1-b_2)b_2\geqslant 0, \quad (5)$$

où on a posé $\xi=(\beta_{12}-\gamma)$ supposé positif, puisque $\gamma>\beta_{12}$ a été traité.

Si $b_1\geqslant a_2+b_2+b_3$, c'est évidemment acquis.

Si $b_1<a_2+b_2+b_3$, soit $\beta_{12}=(b_1-b_2)<a_2+b_3$,

Le discriminant du trinôme du second degré défini par le premier membre de (5) est

$$\left[(b_1-b_2)-(a_2+b_3)\right]^2-4(b_1-b_2)b_2. \quad (6)$$

Il se met sous la forme :

$$\left[(b_1-b_2)-(a_2+2b_2)-b_3)\right]^2+(a_2+b_3)^2$$

$$-(a_2+2b_2+b_3)^2$$

$$=\left[(b_1-b_2)-(a_2+2b_2+b_3)\right]^2$$

$$-4b_2(a_2+b_2+b_3)$$

Compte tenu de ce que le contenu du crochet est négatif, une condition suffisante pour avoir (5) est fournie par

$$(a_2+2b_2+b_3)-(b_1-b_2)\leqslant 2\sqrt{b_2(a_2+b_2+b_3)},$$

soit

$$(b_1-b_2)\geqslant 2\left\{\frac{1}{2}\left[b_2+(a_2+b_2+b_3)\right]-\sqrt{b_2(a_2+b_2+b_3)}\right\} \quad . \quad (7)$$

C'est ce qu'il fallait démontrer.

Le type de résultat précédent veut faire croire que c'est le déchargement de b_1 au niveau de la première ligne qui conduit "plus naturellement" vers la solution optimale.

Ce résultat se généralise en remplaçant a_1,b_1,b_2 et b_3, où $a_1>b_1>b_2$ et où b_3 est le plus grand b_j encore disponible, par a_{i_1}, b_{j_1}, b_{j_2} et b_{j_3}, où $a_{i_1}>b_{j_1}>b_{j_2}$ et où b_{j_3} (qui remplace b_3) est le plus grand b_j encore disponible. On peut aussi intervertir les rôles respectifs des deux marges.

IV. L'ALGORITHME PROGRAMME

IV.1. Introduction

Nous avons déjà exprimé au paragraphe III.1. l'idée générale de l'algorithme et son application à un exemple.

Cet algorithme repose sur les propriétés suivantes :

(i) Si une même part marginale se retrouve : $a_{i_o}=b_{j_o}$, la configuration optimale passe nécessairement par la résolution du couple (a_{i_o},b_{j_o}) (cf.§III.2.).

(ii) S'il existe un seul couple extrémal (a_{i_1},b_{j_1}) pour le préordre total intersection des deux préordres totaux ω_d et ω_s [cf. les expressions logiques **(6)** et **(7)** du paragraphe III.1.], alors la configuration optimale comprend nécessairement la résolution du couple (a_{i_1},b_{j_1}) (cf. §III.3.).

(iii) S'il y a plus d'un seul couple extrémal, la configuration optimale suppose nécessairement la résolution de l'un des couples extrémaux.

Il y a aussi bien sûr la propriété implicite -sans laquelle cette recherche n'aurait été possible- que nous avons déjà mentionnée dans l'introduction : la suppression d'une ligne (resp. colonne) et le réaménagement conséquent de la marge colonne (resp. ligne) préserve le caractère optimal de la configuration obtenue.

Les propriétés **(i)** et **(ii)** ont été respectivement démontrées aux paragraphes III.2. et III.3., celle **(iii)** reste encore à l'état de conjecture que nous avons cherché à étayer tant soit peu au niveau du paragraphe III.4.

Dans le cas général, on commence par ordonner l'ensemble des couples extrémaux conformément à ω_d de la plus petite différence à la plus grande. Ainsi si $(a_{i_o},b_{j_o}),(a_{i_1},b_{j_1}),\ldots,(a_{i_e},b_{j_e})$ est la suite ordonnée des couples extrémaux, on a

$$|a_{i_o}-b_{j_o}|\leqslant|a_{i_1}-b_{j_1}|\leqslant\ldots\leqslant|a_{i_e}-b_{j_e}|$$

et l'algorithme commence par la résolution de (a_{i_o},b_{j_o}).

La portée de l'algorithme aurait sensiblement diminué s'il s'agissait après chaque résolution, d'empiler l'ensemble de tous les couples extrémaux. En réalité, comme le montrera l'important théorème suivant, il suffira d'empiler à partir de (a_{i_o},b_{j_o}), le sous-ensemble de couples extrémaux pour lesquels ou bien l'adresse de la première composante est i_o, ou bien l'adresse de la seconde composante est j_o. En effet, et c'est plus précisément l'objet du théorème qui suit, le caractère extrémal d'un couple (a_{i_1},b_{j_1}) est préservé après la résolution de (a_{i_o},b_{j_o}) si $i_1 \neq i_o$ et $j_1 \neq j_o$.

IV.2. Théorème limitant l'empilement

THEOREME 1. Si (a_{i_o},b_{j_o}) est le couple extrémal qui réalise la plus petite valeur de $|a_i-b_j|$, le déchargement de a_{i_o} dans j_o (si $a_{i_o} \leq b_{j_o}$) [resp. de b_{j_o} dans i_o (si $a_{i_o} \geq b_{j_o}$)] préserve le caractère extrémal de tout autre couple (a_{i_1},b_{i_1}).

On suppose -sans restreindre la généralité- que $a_{i_o} > b_{j_o}$. Deux cas sont à considérer :

1 $a_{i_1} < b_{j_1}$ et 2 $a_{i_1} > b_{j_1}$.

On posera $\alpha_{i_o j_o} = (a_{i_o} - b_{j_o})$ et on a

$\alpha_{i_o j_o} < |a_{i_1} - b_{j_1}|$ et $(a_{i_o} + b_{j_o}) < (a_{i_1} + b_{j_1})$. (1)

Etude du 1er cas 1. On a

$0 < (a_{i_o} - b_{j_o}) < (b_{j_1} - a_{i_1})$ et $(a_{i_o} + b_{j_o}) < (a_{i_1} + b_{j_1})$. (2)

La question se pose de savoir s'il peut exister j tel que $(\alpha_{i_o j_o}, b_j)$ puisse strictement dépasser (a_{i_1},b_{j_1}). Si c'est le cas, on doit avoir

$$\left.\begin{array}{l} |b_j - \alpha_{i_o j_o}| < b_{j_1} - a_{i_1} \\ \text{et} \\ b_j + \alpha_{i_o j_o} > b_{j_1} + a_{i_1} \end{array}\right. \qquad (3)$$

Ici encore, on doit considérer deux situations ; la première est celle où $b_j < \alpha_{i_o j_o}$. Dans ce cas, on aurait

$$\begin{array}{l} \alpha_{i_o j_o} < b_j + b_{j_1} - a_{i_1} \\ \text{et} \\ \alpha_{i_o j_o} > -b_j + b_{j_1} + a_{i_1} \end{array}$$

D'où on tire, en mettant en évidence b_j:

$b_j > a_{i_1}$. (4)

Ce qui entraînerait -compte tenu de $b_j < \alpha_{i_o j_o}$-

$\alpha_{i_o j_o} > a_{i_1}$ (5)

Considérons à présent la deuxième situation où $b_j > \alpha_{i_o j_o}$. Dans ce cas on aurait

$$\left.\begin{array}{l} b_j - \alpha_{i_o j_o} < b_{j_1} - a_{i_1} \\ \text{et} \\ b_j + \alpha_{i_o j_o} > b_{j_1} + a_{i_1} \end{array}\right\} \quad (6)$$

Ce qui implique, en utilisant la transitivité à partir des deux bornes de b_j:

$-a_{i_o} + b_{j_o} + b_{j_1} + a_{i_1} < a_{i_o} - b_{j_o} + b_{j_1} - a_{i_1}$

Il en résulte que

$a_{i_1} < \alpha_{i_o j_o}$.

Donc, les relations (3) impliquent de toute façon la relation (5).

Maintenant, peut-on avoir la relation (5) ?

Si oui, on a a fortiori,

$a_{i_o} > a_{i_1}$ et donc $(a_{i_o} + b_{j_1}) > (a_{i_1} + b_{j_1})$. (7)

D'autre part (cf.(1)),

$a_{i_1} + b_{j_1} > a_{i_o} + b_{j_o}$,

ce qui donne

$b_{j_1} > a_{i_o} - a_{i_1} + b_{j_o}$. (8)

Si on avait $b_{j_o} < a_{i_1}$, on aurait (en vertu de $a_{i_o} > a_{i_1}$),

$a_{i_1} - b_{j_o} < a_{i_o} - b_{j_o}$

et (a_{i_1},b_{j_o}) aurait été considéré avant (a_{i_o},b_{j_o}). Par conséquent,

$$b_{j_o} \geq a_{i_1}$$

Il en résulte à partir de **(8)**, que

$$b_{j_1} > a_{i_o}$$

et toujours en vertu de $a_{i_o} > a_{i_1}$,

$$(b_{j_1}-a_{i_o})<(b_{j_1}-a_{i_1}), \quad \textbf{(9)}$$

ce qui, en considérant **(7)**, conduit à une absurdité : (a_{i_1},b_{j_1}) n'est plus un élément extrémal puisqu'il se trouve dépassé par (a_{i_o},b_{j_1}).

Ainsi, les relations **(3)** qui conduisent à la relation **(5)** sont impossibles.

<u>**Etude du 2ème cas 2**</u> On a

$$0<(a_{i_o}-b_{j_o})<(a_{i_1}-b_{j_1}) \text{ et } (a_{i_o}+b_{j_o})<(a_{i_1}+b_{j_1})$$

Nous allons reprendre la même démarche que ci-dessus. La question se pose de savoir s'il peut exister un indice j pour lequel

$$\left.\begin{array}{l} \alpha_{i_oj_o}+b_j > a_{i_1}+b_{j_1} \\ \text{et} \quad |b_j-\alpha_{i_oj_o}| < a_{i_1}-b_{j_1} \end{array}\right\} \textbf{(10)}$$

Deux situations sont à envisager :

A: $b_j < \alpha_{i_oj_o}$ et B: $b_j > \alpha_{i_oj_o}$

Commençons par considérer la situation A. Les inégalités **(10)** deviennent

$$\alpha_{i_oj_o} > a_{i_1}+b_{j_1}-b_j$$

$$\alpha_{i_oj_o} < a_{i_1}-b_{j_1}+b_j$$

ou encore

$$\left.\begin{array}{l} b_j > -a_{i_o}+b_{j_o}+a_{i_1}+b_{j_1} \\ b_j > a_{i_o}-b_{j_o}-a_{i_1}+b_{j_1} \end{array}\right\} \textbf{(11)}$$

qui par sommation membre à membre donnent

$b_j > b_{j_1}$ **(12)**

La première des deux inégalités **(11)** donne

$b_j + a_{i_0} > a_{i_1} + b_{j_1}$. **(13)**

D'autre part, on a nécessairement

$a_{i_0} < a_{i_1}$, **(14)**

car, autrement, compte tenu du cas 2 où on se trouve : $a_{i_0} > b_{i_0}$, $a_{i_1} > b_{i_1}$ et $(a_{i_0} - b_{i_0}) < (a_{i_1} - b_{i_1})$, on aurait eu

$(a_{i_0} + b_{i_0}) > (a_{i_1} + b_{i_1})$

et (a_{i_1}, b_{i_1}) n'aurait pas été extrémal.

Dans la situation A, on a a fortiori

$b_j < a_{i_0}$.

D'autre part, en vertu de **(12)** et de **(14)**,

$a_{i_0} - b_j < a_{i_0} - b_{j_1} < a_{i_1} - b_{j_1}$

Donc, (a_{i_1}, b_{j_1}) ne serait plus extrémal face à (a_{i_0}, b_j). Donc la situation A est impossible.

Considérons à présent la situation B. Les relations **(10)** deviennent :

$$\left.\begin{array}{l} b_j > (a_{i_1} + b_{j_1}) - (a_{i_0} - b_{j_0}) \\ b_j < (a_{i_1} - b_{j_1}) + (a_{i_0} - b_{j_0}) \end{array}\right. \quad \textbf{(15)}$$

La question se pose de savoir s'il peut exister un tel b_j.

A partir de la première relation (**15**) on a

$(a_{i_1}+b_j)>(a_{i_1}+b_{j_1})+a_{i_1}-(a_{i_o}-b_{j_o})>(a_{i_1}+b_{j_1})$, (**16**)

compte tenu de (**14**).

En orientant de la même manière les deux relations (**15**) et en additionnant terme à terme, on obtient

$\alpha_{i_o j_o} > b_{j_1}$. (**17**)

Par conséquent, compte tenu du cas B où on se trouve,

$b_j > b_{j_1}$. (**18**)

Dans ces conditions, si a_{i_1} (qui est supérieur à b_{j_1}) est supérieur à b_j, on a trivialement

$(a_{i_1}-b_j)<(a_{i_1}-b_{j_1})$ (**19**)

Cette relation alliée à (**16**) montre l'impossibilité. En effet, devant (a_{i_1},b_j), (a_{i_1},b_{j_1}) ne peut être extrémal.

Enfin, si $a_{i_1}<b_j$, la deuxième relation (**15**) donne

$(b_j-a_{i_1})<\alpha_{i_o j_o}-b_{j_1}<\alpha_{i_o j_o}<(a_{i_1}-b_{j_1})$ (**20**)

et l'impossibilité apparaît également pour la même raison. C.Q.F.D.

IV.3. Programmation de l'algorithme

Le sous-programme TRAIT gère les sous-programmes destinés à produire les différents tableaux de contingence et assure la sélection de la meilleure configuration (maximisant la somme des carrés des parts).

La réalisation naturelle de ce sous-programme conduisant à la construction d'un algorithme récursif, il a fallu lui appliquer les techniques classiques d'élimination de la récursivité afin de pouvoir le programmer en FORTRAN. Pour ce faire, une dimension supplémentaire correspondante au niveau de récursivité a dûe être ajoutée aux tableaux des marges, de même que la présence des trois piles IPILLG(*), IPILCL(*) et IPILNV(*) a été rendue nécessaire. Ces trois piles sont destinées à recevoir respectivement, à chaque simulation d'un appel récursif, la ligne, la colonne et le niveau de récursivité de chacune des possibilités qui devront être explorées par la suite.

a) Algorithme

```
BEGIN
    NIV:=0; NPIL:=0;
    co NIV: niveau de récursivité fco
    co NPIL: indice de pile fco
    REPEAT
        FINI:=FALSE;
        CALL DETER (NPIL,FINI)
        DO WHILE (NOT FINI);
            co sauvegarde de l'état des tableaux à ce niveau de
               récursivité fco
            CALL COPIE
            co On continue en exploitant un des cas à envisager fco
            NIV:=NIV+1;
            CALL MAJ;
            CALL DETER (NPIL,FINI)
        END DO
        co On vient d'obtenir un des tableaux possible, on le sélectionne
           s'il est meilleur fco
        IF <meilleur> THEN <sélection> ENDIF
        co Dépilement éventuel d'une possibilité mise en attente fco
        IF (NPIL≠0) THEN
           LIGNE_A_TRAITER:=IPILLG(NPIL);
           COLONNE_A_TRAITER:=IPILCL(NPIL);
           NIV:=IPILNV(NPIL);NPIL:=NPIL-1;
           CALL COPIE;
           NIV:=NIV+1;CALL MAJ
        END IF
    UNTIL (NPIL=0);
END
```

b) Sous-programme DETER :

Son rôle consiste à traiter immédiatement tous les cas déterministes (marge d'une ligne égale à celle d'une colonne, plus petite différence et plus grande somme réalisée simultanément) puis, lorsque plusieurs possibilités différentes sont à explorer, ce sous-programme en conserve une en paramètre après avoir empilé les autres dans IPILLG, IPIL et IPILNC et avant de rendre la main à TRAIT.

c) Sous-programme COPIE :

Il assure la sauvegarde des marges pour un niveau donné.

d) Sous-programme MAJ :

Il est destiné à mettre à jour la table et les marges après le choix d'un cas à tester (prise en compte de ce choix).

```
C
C***********************************************************************
C***********************************************************************
C**                                                                   **
C**                PROGRAMME DE GENERATION DE TABLEAUX                **
C**                ===================================                **
C**                                                                   **
C**           DE CONTINGENCE MAXIMISANT LA SOMME DES CARRES           **
C**           ============================================            **
C**                                                                   **
C**                                                                   **
C**               *** GTAB   VERSION 0.1 (JANVIER 1986) ***           **
C**                                                                   **
C**               AUTEURS : LERMAN I. C.                              **
C**                         (UNIVERSITE RENNES I)                     **
C**                                                                   **
C**                         PETER PH.                                 **
C**                         (UNIVERSITE RENNES I)                     **
C**                                                                   **
C***********************************************************************
C***********************************************************************
C
      PROGRAM GTAB
C
      DIMENSION INTG (30000)
C
      LGINTG=30000
      NUMERR=0
      IFLU=0
      IFEC=0
      IFIC=10
   10 CONTINUE
      WRITE (IFEC,1000)
      READ (IFLU,2000,ERR=10) NL
   20 CONTINUE
      WRITE (IFEC,1010)
      READ (IFLU,2000,ERR=20) NC
C
      NCL=NL*NC
      NCPL=NL+NC
      NCPLS=(NCPL*(NCPL-1))/2
C
C*---------------------------------------------------------------------*
C*         APPEL AU SOUS-PROGRAMME D'ALLOCATION DYNAMIQUE              *
C*---------------------------------------------------------------------*
C
      CALL ALLO (IFEC,LGINTG,NL,NC,NCPL,NCL,NCPLS, IITAB,IITABM,
     1           ILIGNE,IICOLO,INCAST,IIPILL,IIPILC,IIPILN,
     2           IIM,IIP,IILG,IICL,IILGT,IICLT, NUMERR)
C
      IF (NUMERR .EQ. 0) THEN
         OPEN (IFIC)
         REWIND IFIC
```

```
         CALL SPG (IFLU,IFEC,IFIC,NL,NC,NCPL,NCL,NCPLS, INTG(IITAB),
     1             INTG(IITABM),INTG(ILIGNE),INTG(IICOLO),
     2             INTG(INCAST),INTG(IIPILL),INTG(IIPILC),INTG(IIPILN),
     3             INTG(IIM),INTG(IIP),INTG(IILG),INTG(IICL),
     4             INTG(IILGT),INTG(IICLT), NUMERR)
         CLOSE (IFIC)
      END IF
C
C*---------------------------------------------------------------------*
C*          FIN DU PROGRAMME PRINCIPAL                                 *
C*---------------------------------------------------------------------*
C
 1000 FORMAT (//'NOMBRE DE LIGNES ?',$)
 1010 FORMAT (//'NOMBRE DE COLONNES ?',$)
 2000 FORMAT (V)
C
      STOP
      END
```

```
C
C***********************************************************************
C*                                                                     *
C*          SOUS-PROGRAMME D'ALLOCATION DYNAMIQUE                      *
C*                                                                     *
C***********************************************************************
C
      SUBROUTINE ALLO (IFEC,LGINTG,NL,NC,NCPL,NCL,NCPLS, IITAB,
     1                 IITABM,ILIGNE,IICOLO,INCAST,IIPILL,IIPILC,IIPILN,
     2                 IIM,IIP,IILG,IICL,IILGT,IICLT, NUMERR)
C
      NCPL1=NCPL+1
C
C*---------------------------------------------------------------------*
C*          CAS DES TABLEAUX D'ENTIERS                                 *
C*---------------------------------------------------------------------*
C
      IITAB=1
      IITABM=IITAB+NCL
      ILIGNE=IITABM+NCL
      IICOLO=ILIGNE+NL*NCPL1
      INCAST=IICOLO+NC*NCPL1
      IIPILL=INCAST+NCPL1
      IIPILC=IIPILL+NCPLS
      IIPILN=IIPILC+NCPLS
      IIM=IIPILN+NCPLS
      IIP=IIM+NCPL
      IILG=IIP+NCPL
      IICL=IILG+NCPL
      IILGT=IICL+NCPL
      IICLT=IILGT+2*NL
      IFIN=IICLT+2*NC-1
C
C*---------------------------------------------------------------------*
C*          CONTROLE DE LA PLACE DISPONIBLE                            *
C*---------------------------------------------------------------------*
C
      IF (IFIN .GT. LGINTG) THEN
         NUMERR=1
         CALL ERROR (IFEC,NUMERR,IFIN)
      END IF
C
C*---------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'ALLO'                               *
C*---------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C***********************************************************************
C*                                                                     *
C*          SOUS-PROGRAMMME DE TRAITEMENT DES ERREURS                  *
C*                                                                     *
C***********************************************************************
C
      SUBROUTINE ERROR (IFICH,NUM,IFIN)
C
      WRITE (IFICH,1000)
      GO TO (10,20), NUM
C
   10 CONTINUE
      WRITE (IFICH,1010) IFIN
      GO TO 30
C
   20 CONTINUE
      WRITE (IFICH,1020)
      GO TO 30
C
   30 CONTINUE
C
C*---------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'ERROR'                              *
C*---------------------------------------------------------------------*
C
 1000 FORMAT (///,T40,'ERREUR',///)
 1010 FORMAT ('PORTER LA DIMENSION DU TABLEAU "INTG" A AU MOINS :',I7)
 1020 FORMAT ('LA SOMME DES LIGNES EST DIFFERENTE DE CELLE DES',
     1        ' COLONNES')
C
      RETURN
      END
```

```
C
C**********************************************************************
C*                                                                    *
C*         SOUS-PROGRAMME PRINCIPAL                                   *
C*                                                                    *
C**********************************************************************
C
      SUBROUTINE SPG (IFLU,IFEC,IFIC,NL,NC,NCPL,NCL,NCPLS, ITAB,ITABM,
     1                LIGNE,ICOLO,NCASTR,IPILLG,IPILCL,IPILNV,IMOINS,
     2                IPLUS,ILIG,ICOL,ILGTR,ICLTR, NUMERR)
C
      DIMENSION ITAB(NL,NC), ITABM(NL,NC), LIGNE(NL,0:NCPL)
      DIMENSION ICOLO(NC,0:NCPL)
      DIMENSION NCASTR(0:NCPL), IPILLG(NCPLS), IPILNV(NCPLS)
      DIMENSION IMOINS(NCPL), IPLUS(NCPL), ILIG(NCPL)
      DIMENSION ICOL(NCPL), IPILCL(NCPLS)
      DIMENSION ILGTR(NL,2), ICLTR(NC,2)
C
C*--------------------------------------------------------------------*
C*         APPEL AU SOUS-PROGRAMME DE LECTURE/ECRITURE DES MARGES     *
C*--------------------------------------------------------------------*
C
      CALL LIREC (IFLU,IFEC,IFIC,NL,NC,NCPL, LIGNE,ICOLO, NUMERR)
      IF (NUMERR .NE. 0) RETURN
C
C*--------------------------------------------------------------------*
C*         INITIALISATIONS                                            *
C*--------------------------------------------------------------------*
C
      NIV=0
      NCASTR(0)=0
      NPIL=0
      ISOM2=0
      ITOT=0
C
C*--------------------------------------------------------------------*
C*         FIN DES INITIALISATIONS, APPEL AU SOUS-PROGRAMME DE        *
C*         REMPLISSAGE DES TABLEAUX                                   *
C*--------------------------------------------------------------------*
C
      CALL TRAIT (NL,NC,NCPL,NCL,NCPLS,NIV,NPIL, ITAB,ITABM,LIGNE,ICOLO,
     1            NCASTR,IPILLG,IPILCL,IPILNV, ISOM2,ITOT,
     2            IMOINS,IPLUS,ILIG,ICOL,ILGTR,ICLTR)
C
C*--------------------------------------------------------------------*
C*         EDITION DE LA MEILLEURE TABLE                              *
C*--------------------------------------------------------------------*
C
      WRITE (IFEC,1000) ITOT
      WRITE (IFEC,1010) ISOM2
      WRITE (IFIC,1000) ITOT
      WRITE (IFIC,1010) ISOM2
C
```

```
      DO 10 I=1, NL
         WRITE (IFEC,1020) (ITABM(I,J), J=1, NC)
         WRITE (IFIC,1030) (ITABM(I,J), J=1, NC)
   10 CONTINUE
C
C*-------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'SPG'                               *
C*-------------------------------------------------------------------*
C
 1000 FORMAT (////'MEILLEURE DES ',I7,' TABLES TROUVEES AVEC')
 1010 FORMAT ('UN SCORE DE :',I10,///)
 1020 FORMAT (16I5)
 1030 FORMAT (26I5)
C
      RETURN
      END
```

```
C
C*********************************************************************
C*                                                                   *
C*        SOUS-PROGRAMME DE LECTURE/ECRITURE DES MARGES              *
C*                                                                   *
C*********************************************************************
C
      SUBROUTINE LIREC (IFLU,IFEC,IFIC,NL,NC,NCPL, LIGNE,ICOLO, NUMERR)
C
      DIMENSION LIGNE(NL,0:NCPL), ICOLO(NC,0:NCPL)
C
      IS1=0
      IS2=0
      IBL=0
      IBC=0
      CNI2=0.
      CNJ2=0.
      DO 20 I=1, NL
   10    CONTINUE
         WRITE (IFEC,1000) I
         READ (IFLU,2000,ERR=10) M
         LIGNE(I,0)=M
         IS1=IS1+M
         IBL=IBL+M*M
         CNI2=CNI2 + M*(M-1)/2
   20 CONTINUE
C
      WRITE (IFIC,1010)
      WRITE (IFIC,1020) (LIGNE(I,0), I=1, NL)
C
      DO 40 I=1, NC
   30    CONTINUE
         WRITE (IFEC,1030) I
         READ (IFLU,2000,ERR=30) M
         ICOLO(I,0)=M
         IS2=IS2+M
         IBC=IBC+M*M
         CNJ2=CNJ2 + M*(M-1)/2
   40 CONTINUE
C
      WRITE (IFIC,1040)
      WRITE (IFIC,1020) (ICOLO(I,0), I=1, NC)
C
      IF (IS2 .NE. IS1) THEN
         NUMERR=2
         CALL ERROR (IFEC,NUMERR,NUMERR)
      ELSE
         IB=MIN0(IBL,IBC)
         CN2=IS1*(IS1-1)/2
         BORNE=2.*(CNI2*CNJ2/CN2 + (CNI2*(1.-CNI2/CN2)*CNJ2*
     1             (1.-CNJ2/CN2))**0.5) + IS1
         WRITE (IFEC,1050) IB, BORNE
         WRITE (IFIC,1050) IB, BORNE
      END IF
```

```
C
C*-----------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'LIREC'                                *
C*-----------------------------------------------------------------------*
C
 1000 FORMAT (/'SOMME DE LA LIGNE',I5,'?',$)
 1010 FORMAT (///,T40,'MARGES DES LIGNES :',//)
 1020 FORMAT (16I5)
 1030 FORMAT (/'SOMME DE LA COLONNE',I5,' ?',$)
 1040 FORMAT (/////,T40,'MARGES DES COLONNES :',//)
 1050 FORMAT (////,' LA BORNE CALCULEE EST :',I10,4X,
     1        ' LA BORNE ANALYTIQUE EST : ',F11.0,///)
 2000 FORMAT (V)
C
      RETURN
      END
```

```
C
C*********************************************************************
C*                                                                   *
C*          SOUS-PROGRAMME DE REMPLISSAGE DES TABLEAUX               *
C*          (SIMULATION D'UN SOUS-PROGRAMME RECURSIF)                *
C*                                                                   *
C*********************************************************************
C
      SUBROUTINE TRAIT (NL,NC,NCPL,NCL,NCPLS,NIV,NPIL, ITAB,ITABM,LIGNE,
     1                  ICOLO,NCASTR,IPILLG,IPILCL,IPILNV,
     2                   ISOM2,ITOT, IMOINS,IPLUS,ILIG,ICOL,ILGTR,ICLTR)
C
      LOGICAL FINI
C
      DIMENSION ITAB(NL,NC), ITABM(NL,NC), LIGNE(NL,0:NCPL)
      DIMENSION ICOLO(NC,0:NCPL)
      DIMENSION NCASTR(0:NCPL), IPILLG(NCPLS), IPILNV(NCPLS)
      DIMENSION IMOINS(NCPL), IPLUS(NCPL), ILIG(NCPL)
      DIMENSION ICOL(NCPL), IPILCL(NCPLS)
      DIMENSION ILGTR(NL,2), ICLTR(NC,2)
C
   10 CONTINUE
C
C*------------------------------------------------------------------*
C*          RETOUR RECURSIVITE                                      *
C*------------------------------------------------------------------*
C
      FINI=.FALSE.
      CALL DETER (NL,NC,NCPL,NCL,NCPLS,NIV, NPIL,ITAB,LIGNE,ICOLO,
     1           NCASTR,IPILLG,IPILCL,IPILNV,ILELU,ICELU,
     2           FINI, IMOINS,IPLUS,ILIG,ICOL,ILGTR,ICLTR,ITOT)
C
   20 CONTINUE
C
      IF (.NOT. FINI) THEN
         CALL COPIE (NL,NC,NCPL,NIV, LIGNE,ICOLO,NCASTR)
         NIV=NIV+1
         CALL MAJ (NL,NC,NCPL,NIV,ILELU,ICELU, ITAB,LIGNE,ICOLO,NCASTR)
         CALL DETER (NL,NC,NCPL,NCL,NCPLS,NIV, NPIL,ITAB,LIGNE,ICOLO,
     1              NCASTR,IPILLG,IPILCL,IPILNV,ILELU,ICELU,FINI,
     2               IMOINS,IPLUS,ILIG,ICOL,ILGTR,ICLTR,ITOT)
C
C*------------------------------------------------------------------*
C*          SIMULATION APPEL RECURSIF                               *
C*------------------------------------------------------------------*
C
         GO TO 20
      END IF
C
C*------------------------------------------------------------------*
C*          ON REGARDE SI LE TABLEAU DE CONTINGENCE OBTENU          *
C*          EST MEILLEUR QUE LE PRECEDENT                           *
C*------------------------------------------------------------------*
C
```

```
      ISOMME=0
      ITOT=ITOT+1
      DO 40 I=1, NL
         DO 30 J=1, NC
            ISOMME=ISOMME+ITAB(I,J)**2
   30    CONTINUE
   40 CONTINUE
C
      IF (ISOMME .GT. ISOM2) THEN
         ISOM2=ISOMME
         DO 60 I=1, NL
            DO 50 J=1, NC
               ITABM(I,J)=ITAB(I,J)
   50       CONTINUE
   60    CONTINUE
      END IF
C
C*------------------------------------------------------------------------*
C*          ON REVIENT AU NIVEAU PRECEDANT DANS L'ARBRE DE RECURSIVITE *
C*------------------------------------------------------------------------*
C
      IF (NPIL .GT. 0) THEN
         IL=IPILLG(NPIL)
         IC=IPILCL(NPIL)
         NIV=IPILNV(NPIL)
         NPIL=NPIL-1
         CALL COPIE (NL,NC,NCPL,NIV, LIGNE,ICOLO,NCASTR)
         NIV=NIV+1
         CALL MAJ (NL,NC,NCPL,NIV,IL,IC, ITAB,LIGNE,ICOLO,NCASTR)
         GO TO 10
      END IF
C
C*------------------------------------------------------------------------*
C*          SI ON NE PEUT PAS : ON S'ARRETE                                *
C*                                                                          *
C*          FIN DU SOUS-PROGRAMME 'TRAIT'                                  *
C*------------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C*********************************************************************
C*                                                                   *
C*         SOUS-PROGRAMME TRAITANT LES CAS NON RECURSIFS             *
C*                                                                   *
C*********************************************************************
C
      SUBROUTINE DETER (NL,NC,NCPL,NCL,NCPLS,NIV, NPIL,ITAB,LIGNE,ICOLO,
     1                  NCASTR,IPILLG,IPILCL,IPILNV,ILELU,ICELU,FINI,
     2                   IMOINS,IPLUS,ILIG,ICOL,ILGTR,ICLTR,ITOT)
C
      LOGICAL FINI, TRV
      LOGICAL RECUR
C
      DIMENSION ITAB(NL,NC), LIGNE(NL,0:NCPL), ICOLO(NC,0:NCPL)
      DIMENSION NCASTR(0:NCPL)
      DIMENSION IPILLG(NCPLS), IPILNV(NCPLS), IMOINS(NCPL), IPLUS(NCPL)
      DIMENSION ILIG(NCPL), ILGTR(NL,2), ICLTR(NC,2)
      DIMENSION ICOL(NCPL), IPILCL(NCPLS)
C
   10 CONTINUE
C
C*-------------------------------------------------------------------*
C*         TRAITEMENT DES 'EGALITES'                                 *
C*-------------------------------------------------------------------*
C
      DO 60 I=1, NL
         IF (LIGNE(I,NIV) .EQ. 0) GO TO 60
         LIGI=LIGNE(I,NIV)
         J=1
         TRV=.FALSE.
C
C%       DO WHILE (^TRV & J<=NC)
C
   20    CONTINUE
         IF ((TRV) .OR. (J .GT. NC)) GO TO 50
            IF (ICOLO(J,NIV) .EQ. LIGI) THEN
               TRV=.TRUE.
               NCAS=NCASTR(NIV)
               DO 30 I1=1, NL
                  IF (LIGNE(I1,NIV) .NE. 0) THEN
                     ITAB(I1,J)=0
                     NCAS=NCAS+1
                  END IF
   30          CONTINUE
C
               DO 40 J1=1, NC
                  IF (ICOLO(J1,NIV) .NE. 0) THEN
                     ITAB(I,J1)=0
                     NCAS=NCAS+1
                  END IF
   40          CONTINUE
               ITAB(I,J)=LIGI
               LIGNE(I,NIV)=0
```

```
                 ICOLO(J,NIV)=0
                 NCASTR(NIV)=NCAS-1
              END IF
              J=J+1
              GO TO 20
   50      CONTINUE
C
C%         END DO
C
   60 CONTINUE
C
C*-----------------------------------------------------------------------*
C*            ON A TRAITE LES 'EGALITES'                                 *
C*-----------------------------------------------------------------------*
C
      IF (NCASTR(NIV) .EQ. NCL) THEN
         FINI=.TRUE.
         RETURN
      END IF
C
C*-----------------------------------------------------------------------*
C*       CALCUL DES 'SOMMES' ET DES 'DIFFERENCES'                        *
C*-----------------------------------------------------------------------*
C
      CALL TRIM (NL,NCPL,LIGNE,NIV, ILGTR)
      CALL TRIM (NC,NCPL,ICOLO,NIV, ICLTR)
      IT=0
      DO 70 I=1, NL
         IF (LIGNE(I,NIV) .NE. 0) THEN
            MGL=LIGNE(I,NIV)
            J=IDICH (NC,ICLTR,MGL)
            IF (J .GT. 0) THEN
               IT=IT+1
               ILIG(IT)=I
               ICOL(IT)=J
               IMOINS(IT)=IABS(MGL-ICOLO(J,NIV))
               IPLUS(IT)=MGL+ICOLO(J,NIV)
            END IF
         END IF
   70 CONTINUE
C
      DO 80 J=1, NC
         IF (ICOLO(J,NIV) .NE. 0) THEN
            MGC=ICOLO(J,NIV)
            I=IDICH (NL,ILGTR,MGC)
            IF (I .GT. 0) THEN
               IT=IT+1
               ILIG(IT)=I
               ICOL(IT)=J
               IMOINS(IT)=IABS(MGC-LIGNE(I,NIV))
               IPLUS(IT)=MGC+LIGNE(I,NIV)
            END IF
         END IF
   80 CONTINUE
```

```
C
C*--------------------------------------------------------------------*
C*          TRI SUIVANT LES DIFFERENCES CROISSANTES ET LES SOMMES      *
C*          DECROISSANTES                                              *
C*--------------------------------------------------------------------*
C
      CALL TRITB (NCPL,IT,NL,NC, IMOINS,IPLUS,ILIG,ICOL,ILGTR,ICLTR,
     1            NEL,NEC)
C
C*--------------------------------------------------------------------*
C*       SELECTION DES LIGNES ET DES COLONNES A TRAITER                *
C*--------------------------------------------------------------------*
C
      ILELU=ILIG(1)
      ICELU=ICOL(1)
      IBOMIN=IPLUS(1)
      K=1
      DO 90 I=2, IT
         IPLUSI=IPLUS(I)
         ILII=ILIG(I)
         ICOI=ICOL(I)
         IF (RECUR (ILELU,ICELU,ILII,ICOI,IPLUSI, NEL,NEC,IBOMIN)) THEN
            K=K+1
            IBOMIN=IPLUSI
            NPIL=NPIL+1
            IPILLG(NPIL)=ILII
            IPILCL(NPIL)=ICOI
            IPILNV(NPIL)=NIV
         END IF
   90 CONTINUE
C
C*--------------------------------------------------------------------*
C*          SI LA PLUS GRANDE SOMME ET LA PLUS PETITE DIFFERENCE       *
C*          SONT REALISEES AU MEME ENDROIT ON CONTINUE                 *
C*          SINON ON S'ARRETE                                          *
C*--------------------------------------------------------------------*
C
      IF (K .GT. 1) RETURN
C
      IL=ILELU
      IC=ICELU
      CALL MAJ (NL,NC,NCPL,NIV,IL,IC, ITAB,LIGNE,ICOLO,NCASTR)
      IF (NCASTR(NIV) .EQ. NCL) THEN
         FINI=.TRUE.
      ELSE
         GO TO 10
      END IF
C
C*--------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'DETER'                              *
C*--------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C**********************************************************************
C*                                                                    *
C*          SOUS-PROGRAMME DE TRI DES MARGES                          *
C*                                                                    *
C**********************************************************************
C
      SUBROUTINE TRIM (N,M,MARG,NV, MGTR)
C
      DIMENSION MARG(N,0:M), MGTR(N,2)
C
      DO 10 I=1, N
         MGTR(I,1)=MARG(I,NV)
         MGTR(I,2)=I
   10 CONTINUE
C
C*--------------------------------------------------------------------*
C*          TRI DE SHELL-METZNER                                      *
C*--------------------------------------------------------------------*
C
      IPAS=N/2
   20 K=N-IPAS
      DO 40 J=1, K
         I=J
   30    L=I+IPAS
         IF (MGTR(I,1) .GT. MGTR(L,1)) THEN
            ITAMP=MGTR(I,1)
            MGTR(I,1)=MGTR(L,1)
            MGTR(L,1)=ITAMP
            ITAMP=MGTR(I,2)
            MGTR(I,2)=MGTR(L,2)
            MGTR(L,2)=ITAMP
            I=I-IPAS
            IF (I .GE. 1) GO TO 30
         END IF
   40 CONTINUE
      IPAS=IPAS/2
      IF (IPAS .GE. 1) GO TO 20
C
C*--------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'TRIM'                              *
C*--------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C*********************************************************************
C*                                                                   *
C*          FONCTION DE RECHERCHE DICHOTOMIQUE                       *
C*                                                                   *
C*********************************************************************
C
      FUNCTION IDICH (N,MARG,M)
C
      DIMENSION MARG(N,2)
C
      IBI=1
      IBS=N
C
   10 CONTINUE
      IBM=(IBI+IBS)/2
      IF (MARG(IBM,1) .GT. M) THEN
         IBS=IBM-1
         IF (IBS .GE. IBI) GOTO 10
         IF (IBS .EQ. 0) THEN
            IDICH=0
            RETURN
         END IF
         IND=MARG(IBS,2)
         IF (MARG(IBS,1) .EQ. 0) IND=0
         IDICH=IND
         RETURN
      ELSE IF (MARG(IBM,1) .LT. M) THEN
         IBI=IBM+1
         IF (IBI .LE. IBS) GOTO 10
         IND=MARG(IBM,2)
         IF (MARG(IBM,1) .EQ. 0) IND=0
         IDICH=IND
         RETURN
      ELSE
         IDICH=MARG(IBM,2)
         RETURN
      END IF
C
      IDICH=0
C
C*-------------------------------------------------------------------*
C*          FIN DE LA FONCTION 'IDICH'                               *
C*-------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C***********************************************************************
C*                                                                     *
C*          SOUS-PROGRAMME DE TRI SUIVANT LES DIFFERENCES CROISSANTES  *
C*          ET LES SOMMES DECROISSANTES                                *
C*                                                                     *
C***********************************************************************
C
      SUBROUTINE TRITB (NCPL,N,NL,NC, IMOINS,IPLUS,ILIG,ICOL,ILGTR,
     1                  ICLTR,NEL,NEC)
C
      DIMENSION IMOINS(NCPL), IPLUS(NCPL), ILIG(NCPL), ICOL(NCPL)
      DIMENSION ILGTR(NL,2), ICLTR(NC,2)
C
      IPAS=N/2
   10 K=N-IPAS
      DO 30 J=1, K
         I=J
   20    L=I+IPAS
         IF ((IMOINS(I) .GT. IMOINS(L)) .OR. ((IMOINS(I) .EQ. IMOINS(L))
     1        .AND. (IPLUS(I) .LT. IPLUS(L)))) THEN
            ITP=IMOINS(I)
            IMOINS(I)=IMOINS(L)
            IMOINS(L)=ITP
            ITP=IPLUS(I)
            IPLUS(I)=IPLUS(L)
            IPLUS(L)=ITP
            ITP=ILIG(I)
            ILIG(I)=ILIG(L)
            ILIG(L)=ITP
            ITP=ICOL(I)
            ICOL(I)=ICOL(L)
            ICOL(L)=ITP
            I=I-IPAS
            IF (I .GE. 1) GO TO 20
         END IF
   30 CONTINUE
      IPAS=IPAS/2
      IF (IPAS .GE. 1) GO TO 10
C
      NEL=0
      J=ILIG(1)
      DO 40 I=1, NL
         IF (ILGTR(I,1) .EQ. J) NEL=NEL+1
   40 CONTINUE
      NEC=0
      J=ICOL(1)
      DO 50 I=1, NC
         IF (ICLTR(I,1) .EQ. J) NEC=NEC+1
   50 CONTINUE
```

```
C
C*-----------------------------------------------------------------*
C*         FIN DU SOUS-PROGRAMME 'TRITB'                           *
C*-----------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C*********************************************************************
C*                                                                   *
C*          FONCTION LOGIQUE PRECISANT SI IL EST NECESSAIRE          *
C*          D'EMPILER LA LIGNE ET LA COLONNE CONSIDEREES             *
C*                                                                   *
C*********************************************************************
C
      LOGICAL FUNCTION RECUR (IL,IC,IL1,IC1,IP1, NEL,NEC,IBN)
C
      IF (((IL .NE. IL1) .AND. (IC .NE. IC1)) .OR. (IP1 .LE. IBN)) THEN
         RECUR=.FALSE.
         RETURN
      ELSE IF (IL .EQ. IL1) THEN
         NEL=NEL-1
         IF (NEL .GE. 1) THEN
            RECUR=.FALSE.
         ELSE
            RECUR=.TRUE.
         END IF
         RETURN
      ELSE
         NEC=NEC-1
         IF (NEC .GE. 1) THEN
            RECUR=.FALSE.
         ELSE
            RECUR=.TRUE.
         END IF
         RETURN
      END IF
C
C*-------------------------------------------------------------------*
C*          FIN DE LA FONCTION 'RECUR'                               *
C*-------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C*********************************************************************
C*                                                                   *
C*        SOUS-PROGRAMME DE MISE A JOUR AVANT ET APRES EMPILEMENT    *
C*                                                                   *
C*********************************************************************
C
      SUBROUTINE MAJ (NL,NC,NCPL,NIV,IL,IC, ITAB,LIGNE,ICOLO,NCASTR)
C
      DIMENSION ITAB(NL,NC), LIGNE(NL,0:NCPL), ICOLO(NC,0:NCPL)
      DIMENSION NCASTR(0:NCPL)
C
      LI=MINO (LIGNE(IL,NIV), ICOLO(IC,NIV))
      NCAS=NCASTR(NIV)
C
      IF (LIGNE(IL,NIV) .LT. ICOLO(IC,NIV)) THEN
         DO 10 J=1, NC
            IF (ICOLO(J,NIV) .NE. 0) THEN
               ITAB(IL,J)=0
               NCAS=NCAS+1
            END IF
   10    CONTINUE
         ICOLO(IC,NIV)=ICOLO(IC,NIV)-LI
         LIGNE(IL,NIV)=0
      ELSE
         DO 20 I=1, NL
            IF (LIGNE(I,NIV) .NE. 0) THEN
               ITAB(I,IC)=0
               NCAS=NCAS+1
            END IF
   20    CONTINUE
         LIGNE(IL,NIV)=LIGNE(IL,NIV)-LI
         ICOLO(IC,NIV)=0
      END IF
      NCASTR(NIV)=NCAS
      ITAB(IL,IC)=LI
C
C*-------------------------------------------------------------------*
C*        FIN DU SOUS-PROGRAMME 'MAJ'                                *
C*-------------------------------------------------------------------*
C
      RETURN
      END
```

```
C
C**********************************************************************
C*                                                                    *
C*          SOUS-PROGRAMME DESTINE A EMPILER LES TABLEAUX LORS        *
C*          D'UNE SIMULATION D'APPEL RECURSIF                         *
C*                                                                    *
C**********************************************************************
C
      SUBROUTINE COPIE (NL,NC,NCPL,NIV, LIGNE,ICOLO,NCASTR)
C
      DIMENSION LIGNE(NL,0:NCPL), ICOLO(NC,0:NCPL), NCASTR(0:NCPL)
C
      NVP1=NIV+1
C
      DO 10 I=1, NL
         LIGNE(I,NVP1)=LIGNE(I,NIV)
   10 CONTINUE
C
      DO 20 I=1, NC
         ICOLO(I,NVP1)=ICOLO(I,NIV)
   20 CONTINUE
C
      NCASTR(NVP1)=NCASTR(NIV)
C
C*-------------------------------------------------------------------*
C*          FIN DU SOUS-PROGRAMME 'COPIE'                            *
C*-------------------------------------------------------------------*
C
      RETURN
      END
```

IV.5. Exemples

MARGES DES LIGNES :

24 12 6 3

MARGES DES COLONNES :

17 7 5 5 4 4 2 1

LA BORNE CALCULEE EST : 425 LA BORNE ANALYTIQUE EST : 558.

MEILLEURE DES 7 TABLES TROUVEES AVEC
UN SCORE DE : 419

17	0	0	0	4	1	2	0
0	7	5	0	0	0	0	0
0	0	0	5	0	0	0	1
0	0	0	0	0	3	0	0

MARGES DES LIGNES :

24 19 18 5 4

MARGES DES COLONNES :

37 12 11 7 3

LA BORNE CALCULEE EST : 1302 LA BORNE ANALYTIQUE EST : 1478.

MEILLEURE DES 306 TABLES TROUVEES AVEC
UN SCORE DE : 998

24	0	0	0	0
0	12	0	7	0
13	0	5	0	0
0	0	5	0	0
0	0	1	0	3

MARGES DES LIGNES :

30 20 10

MARGES DES COLONNES :

25 15 12 8

LA BORNE CALCULEE EST : 1058 LA BORNE ANALYTIQUE EST : 1210.

MEILLEURE DES 3 TABLES TROUVEES AVEC
UN SCORE DE : 988

25	0	0	5
0	15	2	3
0	0	10	0

MARGES DES LIGNES :

800 500 500 100 70 10 10 7 1

MARGES DES COLONNES :

1000 700 100 80 50 50 8 3 3 2 2

LA BORNE CALCULEE EST : 1155150 LA BORNE ANALYTIQUE EST : 1315408.

MEILLEURE DES 3981 TABLES TROUVEES AVEC
UN SCORE DE : 1010088

0	700	0	0	50	50	0	0	0	0	0
500	0	0	0	0	0	0	0	0	0	0
500	0	0	0	0	0	0	0	0	0	0
0	0	100	0	0	0	0	0	0	0	0
0	0	0	70	0	0	0	0	0	0	0
0	0	0	10	0	0	0	0	0	0	0
0	0	0	0	0	0	8	0	0	2	0
0	0	0	0	0	0	0	3	3	0	1
0	0	0	0	0	0	0	0	0	0	1

REFERENCES

L. HUBERT, Ph. ARABIE (1985) ; "Comparing Partitions", Fourth European meeting of the Classification Societies, Cambridge, July 1985.

I.C. LERMAN (1981) ; "Classification et analyse ordinale des données", Dunod, Paris (1981).

F. MARCOTORCHINO (1984) ; "Utilisation des comparaisons par paires en statistique des contingences (Partie I)", Etude F-069, Centre Scientifique IBM-France, Février 1984.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique